

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 18:04:50 ; Search time 15 Seconds

(without alignments)
262.845 Million cell updates/sec

Title: US-09-142-524D-3

Perfect score: 711

Sequence: 1 MKVTVAENQGPNNRVFIKR.....KFIKRVGIIAAYQNPASMK 134

Scoring table: BLOSUM62

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/aa/5A COMB.pep:*
- 2: /cgn2_6/prodata/1/aa/5B COMB.pep:*
- 3: /cgn2_6/prodata/1/aa/6A COMB.pep:*
- 4: /cgn2_6/prodata/1/aa/6B COMB.pep:*
- 5: /cgn2_6/prodata/1/aa/PCITUS COMB.pep:*
- 6: /cgn2_6/prodata/1/aa/backfillseq.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	188	26.4	60	3	US-08-467-023-62
2	188	26.4	374	3	US-08-467-023-2
3	150	21.1	367	3	US-08-467-023-95
4	150	21.1	370	3	US-08-467-023-97
5	148.5	20.9	514	3	US-08-467-023-134
6	115	16.2	128	3	US-08-467-023-187
7	105	14.8	20	3	US-08-467-023-33
8	99	13.9	127	3	US-08-467-023-188
9	98.5	13.9	387	1	US-08-290-448A-72
10	98.5	13.9	387	1	US-08-175-069A-72
11	98.5	13.9	387	1	US-08-461-939B-72
12	98.5	13.9	387	4	US-08-461-939B-72
13	98.5	13.9	387	4	US-08-290-448A-76
14	98.5	13.9	387	4	US-08-464-000-72
15	97	13.6	397	1	US-08-290-448A-76
16	97	13.6	397	1	US-08-175-069A-76
17	97	13.6	397	4	US-08-461-939B-76
18	97	13.6	397	4	US-08-464-000-76
19	93	13.1	391	1	US-08-290-448A-59
20	93	13.1	391	1	US-08-175-069A-59
21	93	13.1	391	1	US-08-461-939B-59
22	93	13.1	391	4	US-08-464-000-59
23	93	13.1	398	1	US-08-290-448A-74
24	93	13.1	398	1	US-08-175-069A-74
25	93	13.1	398	1	US-08-461-939B-74
26	93	13.1	398	4	US-08-464-000-74
27	93	13.1	398	4	US-08-290-448A-74

28	93	13.1	398	4	US-08-464-000-74	Sequence 74, Appl
29	91.5	12.9	30	3	US-08-467-023-71	Sequence 71, Appl
30	91	12.8	20	3	US-08-467-023-34	Sequence 34, Appl
31	88	12.4	383	1	US-08-290-448A-78	Sequence 78, Appl
32	88	12.4	383	1	US-08-290-448A-78	Sequence 78, Appl
33	88	12.4	383	1	US-08-175-069A-78	Sequence 78, Appl
34	88	12.4	383	4	US-08-461-939B-78	Sequence 78, Appl
35	88	12.4	383	4	US-08-464-000-78	Sequence 80, Appl
36	86	12.1	388	1	US-08-290-448A-80	Sequence 80, Appl
37	86	12.1	388	1	US-08-175-069A-80	Sequence 80, Appl
38	86	12.1	388	4	US-08-461-939B-80	Sequence 80, Appl
39	86	12.1	388	4	US-08-464-000-80	Sequence 80, Appl
40	83.5	11.7	30	3	US-08-467-023-67	Sequence 67, Appl
41	81.5	11.5	24	4	US-08-467-023-72	Sequence 72, Appl
42	81.5	11.5	234	4	US-09-219-983A-7	Sequence 7, Appl
43	81.5	11.4	20	3	US-08-467-023-32	Sequence 32, Appl
44	81	11.4	127	3	US-08-467-023-189	Sequence 189, Appl
45	81	11.4				

ALIGNMENTS

RESULT 1
US-08-467-023-62
Sequence 62, Application US/08467023
Patent No. 6090386

GENERAL INFORMATION:

APPLICANT: Griffith, Irwin J.;
APPLICANT: Pollock, Joanne;
APPLICANT: Bond, Julian F.;
APPLICANT: Gorman, Richard D;
APPLICANT: Kuo, Wei-Chang;
APPLICANT: Yeung, Siu-mei H.;
APPLICANT: Brauer, Andrew;
APPLICANT: Exley, Mark A.;
APPLICANT: Powers, Steven P.
TITLE OF INVENTION: Allergenic Proteins And Peptides From Japanese Cedar Pollen
NUMBER OF SEQUENCES: 261

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA

ZIP: 02154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,023
FILING DATE: June 6, 1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/350,225
FILING DATE: December 6, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jane E. Remillard
REGISTRATION NUMBER: 38,872
REFERENCE/DOCKET NUMBER: 025.6 USD2 (INT-028CPD2)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal

US-08-467-023-62

Query Match 26.4%; Score 188; DB 3; Length 60;
Best Local Similarity 71.2%; Pred. No. 1.5e-17;
Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

OY 73 RPLWIFSGNNIKLKMPMYIAGYKTFDGRRAEVSYYHVNAGKFIIRVDGII 124
Db 6 RPLWIFSGNNIKLKMPMYIAGYKTFDGRRAEVSYYHVNAGKFIIRVDGII 57

RESULT 2

US-08-467-023-2

Sequence 2, Application US/08467023

Patent No. 6090386

GENERAL INFORMATION:

APPLICANT: Griffith, Irwin J.;

APPLICANT: Pollock, Joanne;

APPLICANT: Bond, Julian F.;

APPLICANT: Garman, Richard D.;

APPLICANT: Kuo, Mei-Chang;

APPLICANT: Yeung, Siu-mei H.;

APPLICANT: Brauer, Andrew;

APPLICANT: Exley, Mark A.;

APPLICANT: Powers, Steven P.;

TITLE OF INVENTION: Allergenic Proteins And Peptides From

NUMBER OF SEQUENCES: 261

CORRESPONDENCE ADDRESS:

ADDRESSER: Immunologic Pharmaceutical Corporation, Inc.

STREET: 610 Lincoln St

CITY: Waltham

STATE: MA

COUNTRY: USA

ZIP: 02154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467,023

FILING DATE: June 6, 1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/350,225

FILING DATE: December 6, 1994

ATTORNEY/AGENT INFORMATION:

NAME: Jane E. Remillard

REGISTRATION NUMBER: 38,872

REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 374 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-467-023-2

Query Match 26.4%; Score 188; DB 3; Length 374;
Best Local Similarity 71.2%; Pred. No. 1.9e-16;
Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

OY 73 RPLWIFSGNNIKLKMPMYIAGYKTFDGRRAEVSYYHVNAGKFIIRVDGII 124
Db 87 RPLWIFSGNNIKLKMPMYIAGYKTFDGRRAEVSYYHVNAGKFIIRVDGII 138

RESULT 3

US-08-467-023-95

Sequence 95, Application US/08467023

Patent No. 6090386

GENERAL INFORMATION:

APPLICANT: Griffith, Irwin J.;

APPLICANT: Pollock, Joanne;

APPLICANT: Bond, Julian F.;

APPLICANT: Garman, Richard D.;

APPLICANT: Kuo, Mei-Chang;

APPLICANT: Yeung, Siu-mei H.;

APPLICANT: Brauer, Andrew;

APPLICANT: Exley, Mark A.;

APPLICANT: Powers, Steven P.;

TITLE OF INVENTION: Allergenic Proteins And Peptides From

NUMBER OF SEQUENCES: 261

CORRESPONDENCE ADDRESS:

ADDRESSER: Immunologic Pharmaceutical Corporation, Inc.

STREET: 610 Lincoln St

CITY: Waltham

STATE: MA

COUNTRY: USA

ZIP: 02154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467,023

FILING DATE: June 6, 1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/350,225

FILING DATE: December 6, 1994

ATTORNEY/AGENT INFORMATION:

NAME: Jane E. Remillard

REGISTRATION NUMBER: 38,872

REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 95:

SEQUENCE CHARACTERISTICS:

LENGTH: 367 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-467-023-95

Query Match 21.1%; Score 150; DB 3; Length 367;
Best Local Similarity 55.8%; Pred. No. 1.9e-11;
Matches 29; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

OY 73 RPLWIFSGNNIKLKMPMYIAGYKTFDGRRAEVSYYHVNAGKFIIRVDGII 124
Db 87 KALWIFSGNNIKLKMPMYIAGYKTFDGRRAEVSYYHVNAGKFIIRVDGII 138

RESULT 4

US-08-467-023-97

Sequence 97, Application US/08467023

Patent No. 6090386

GENERAL INFORMATION:

APPLICANT: Griffith, Irwin J.;

APPLICANT: Pollock, Joanne;

APPLICANT: Bond, Julian F.;

APPLICANT: Garman, Richard D.;

APPLICANT: Kuo, Mei-Chang;

APPLICANT: Yeung, Siu-mei H.;

APPLICANT: Brauer, Andrew;

APPLICANT: Exley, Mark A.;

APPLICANT: Powers, Steven P.;

TITLE OF INVENTION: Allergenic Proteins And Peptides From

US-08-467-023-95

TITLE OF INVENTION: Japanese Cedar Pollen
NUMBER OF SEQUENCES: 261
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,023
FILING DATE: June 6, 1995
CLASSIFICATION: 424
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER: 08/350,225
FILING DATE: December 6, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jane E. Remillard
REGISTRATION NUMBER: 38,872
REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMT-028CPD2)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 97:
SEQUENCE CHARACTERISTICS:
LENGTH: 370 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-467-023-97

Query Match 21.1%; Score 150; DB 3; Length 370;
Best Local Similarity 55.8%; Pred. No. 1.9e-11;
Matches 29; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 73 RPLMTISGNNITLKMPTIAGYKTPDGRRAEVSYYHVGAKFIRRVDCII 124
DB 87 KALWIFSONNNIKLKMPLVYAGHKTIIDRGADVHLNGGRCPLFMKRVSHVI 138

RESULT 5
US-08-467-023-134
Sequence 134, Application US/08467023
Patent No. 6090386

GENERAL INFORMATION:
APPLICANT: Griffith, Irwin J.;
APPLICANT: Pollock, Joanne;
APPLICANT: Bond, Julian F.;
APPLICANT: Garman, Richard D;
APPLICANT: Kuo, Mei-Chang;
APPLICANT: Yeung, Siu-mei H.;
APPLICANT: Brauer, Andrew;
APPLICANT: Exley, Mark A.;
APPLICANT: Powers, Steven P.
TITLE OF INVENTION: Allergenic Proteins And Peptides From
NUMBER OF SEQUENCES: 261
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 08/350,225
FILING DATE: December 6, 1994
ATTORNEY/AGENT INFORMATION:

SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,023
FILING DATE: June 6, 1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/350,225
FILING DATE: December 6, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jane E. Remillard
REGISTRATION NUMBER: 38,872
REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMT-028CPD2)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 514 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-467-023-134

Query Match 20.9%; Score 148.5; DB 3; Length 514;
Best Local Similarity 41.4%; Pred. No. 4.8e-11;
Matches 41; Conservative 8; Mismatches 21; Indels 29; Gaps 2;

QY 31 IDIFASKNFHLOKNTIGTGRKNNRIMLOFAKLTGFLMGRRPLMTIFSGNNITLKM 90
DB 236 IDIFASKNFHLOKNTIGTGRKNNRIMLOFAKLTGFLMGRRPLMTIFSGNNITLKM 90

QY 91 MYIAGYKTPDGRRAEVSYYHVGAKFIRRVDCII 123
DB 273 ICGPHGISISLGRNSRAEVSYYHVGAKFIRRVDCII 311

RESULT 6
US-08-467-023-187
Sequence 187, Application US/08467023
Patent No. 6090386

GENERAL INFORMATION:
APPLICANT: Griffith, Irwin J.;
APPLICANT: Pollock, Joanne;
APPLICANT: Bond, Julian F.;
APPLICANT: Garman, Richard D;
APPLICANT: Kuo, Mei-Chang;
APPLICANT: Yeung, Siu-mei H.;
APPLICANT: Brauer, Andrew;
APPLICANT: Exley, Mark A.;
APPLICANT: Powers, Steven P.
TITLE OF INVENTION: Allergenic Proteins And Peptides From
NUMBER OF SEQUENCES: 261
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,023
FILING DATE: June 6, 1995
CLASSIFICATION: 424
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER: 08/350,225
FILING DATE: December 6, 1994
ATTORNEY/AGENT INFORMATION:

350 ID NO. 131

NAME: Jane E. Remillard
REGISTRATION NUMBER: 38,872
REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 187:
SEQUENCE CHARACTERISTICS:
LENGTH: 128 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-467-023-187

Query Match 16.2%; Score 115; DB 3; Length 128;
Best Local Similarity 95.2%; Pred. No. 1.8e-07;
Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 52 WKNNRIWQFAKLGFTLMGR 72
DB 98 WKNNRIWQFAKLGFTLMGR 118

RESULT 7

US-08-467-023-33
Sequence 33, Application US/08467023
Patent No. 6090386

GENERAL INFORMATION:
APPLICANT: Griffith, Irwin J.;
APPLICANT: Pollock, Joanne;
APPLICANT: Bond, Julian F.;
APPLICANT: Garman, Richard D;
APPLICANT: Kuo, Mei-Chang;
APPLICANT: Yeung, Siu-mei H.;
APPLICANT: Brauer, Andrew;
APPLICANT: Exley, Mark A.;
APPLICANT: Powers, Steven P.
TITLE OF INVENTION: Allergenic Proteins And Peptides From
NUMBER OF SEQUENCES: 261

CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.

STREET: 610 Lincoln St

CITY: Waltham

STATE: MA

COUNTRY: USA

ZIP: 02154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,023

FILING DATE: June 6, 1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/350,225

FILING DATE: December 6, 1994

ATTORNEY/AGENT INFORMATION:
NAME: Jane E. Remillard

REGISTRATION NUMBER: 38,872

REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: internal
US-08-467-023-33

Query Match 14.8%; Score 105; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.7e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 IFSGNNNITKLKMPNYIAGYK 97
DB 1 IFSGNNNITKLKMPNYIAGYK 20

RESULT 8

US-08-467-023-188
Sequence 188, Application US/08467023
Patent No. 6090386

GENERAL INFORMATION:
APPLICANT: Griffith, Irwin J.;
APPLICANT: Pollock, Joanne;
APPLICANT: Bond, Julian F.;
APPLICANT: Garman, Richard D;
APPLICANT: Kuo, Mei-Chang;
APPLICANT: Yeung, Siu-mei H.;
APPLICANT: Brauer, Andrew;
APPLICANT: Exley, Mark A.;
APPLICANT: Powers, Steven P.
TITLE OF INVENTION: Allergenic Proteins And Peptides From
NUMBER OF SEQUENCES: 261

CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.

STREET: 610 Lincoln St

CITY: Waltham

STATE: MA

COUNTRY: USA

ZIP: 02154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,023

FILING DATE: June 6, 1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/350,225

FILING DATE: December 6, 1994

ATTORNEY/AGENT INFORMATION:
NAME: Jane E. Remillard

REGISTRATION NUMBER: 38,872

REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 188:

SEQUENCE CHARACTERISTICS:

LENGTH: 127 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: internal

US-08-467-023-188

Query Match 13.9%; Score 99; DB 3; Length 127;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 IDIFASKNFHLQNTITG 49
DB 83 IDIFASKNFHLQNTITG 101

RESULT 9
US-08-290-448A-72
Sequence 72, Application US/08290448A
Patent No. 5676954
GENERAL INFORMATION:
APPLICANT: Rogers, Bruce
APPLICANT: Klapper, David G.
APPLICANT: Rafnar, Thorunn
APPLICANT: Kuo, Mei-chang
TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,448A
FILING DATE: August 15, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/529,951
FILING DATE: May 29, 1990
APPLICATION NUMBER: US 07/325,365
FILING DATE: March 17, 1989
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-018CN
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-290-448A-72

Query Match 13.9%; Score 98.5; DB 1; Length 387;
Best Local Similarity 34.7%; Pred. No. 0.00012;
Matches 26; Conservative 12; Mismatches 22; Indels 15; Gaps 3;

QY 73 RPLMIFSGNNMNIKMKMPYIAGYTFDGRRAEVSYVH---VNGAK--FIRVD----- 121
DB 104 RPLMIFERDWMVIRLDKEMVNSDKTIDRGAKVEIINAGFTLNGVKNVLIHINMEDIK 163
QY 122 ---GIIAAYQNPAS 132
DB 164 VNPGLIKSNDGPAA 178

RESULT 10
US-08-290-448A-72
Sequence 72, Application US/08290448A
Patent No. 5698204
GENERAL INFORMATION:
APPLICANT: Rogers, Bruce
APPLICANT: Klapper, David G.
APPLICANT: Rafnar, Thorunn
APPLICANT: Kuo, Mei-chang
TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510

CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,448A
FILING DATE: August 15, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/529,951
FILING DATE: May 29, 1990
APPLICATION NUMBER: US 07/325,365
FILING DATE: March 17, 1989
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-018CN
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-290-448A-72

Query Match 13.9%; Score 98.5; DB 1; Length 387;
Best Local Similarity 34.7%; Pred. No. 0.00012;
Matches 26; Conservative 12; Mismatches 22; Indels 15; Gaps 3;

QY 73 RPLMIFSGNNMNIKMKMPYIAGYTFDGRRAEVSYVH---VNGAK--FIRVD----- 121
DB 104 RPLMIFERDWMVIRLDKEMVNSDKTIDRGAKVEIINAGFTLNGVKNVLIHINMEDIK 163
QY 122 ---GIIAAYQNPAS 132
DB 164 VNPGLIKSNDGPAA 178

RESULT 11
US-08-175-069A-72
Sequence 72, Application US/08175069A
Patent No. 5776761
GENERAL INFORMATION:
APPLICANT: Rogers, Bruce
APPLICANT: Klapper, David G.
APPLICANT: Rafnar, Thorunn
APPLICANT: Kuo, Mei-chang
TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/175,069A
FILING DATE: December 29, 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/529,951

FILED DATE: May 29, 1990
APPLICATION NUMBER: US 07/325,365
FILING DATE: March 17, 1989
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-018DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-175-069A-72

Query Match 13.9%; Score 98.5; DB 1; Length 387;
Best Local Similarity 34.7%; Pred. No. 0.00012;
Matches 26; Conservative 12; Mismatches 22; Indels 15; Gaps 3;

Qy 73 RPLWIFSGNNMIKIKMPYIAGYKTFDGRRAEVSYYH---VNGAK--FIRRD----- 121
Db 104 RPLWIFSGNNMIKIKMPYIAGYKTFDGRRAEVSYYH---VNGAK--FIRRD----- 163
Qy 122 ---GIIAAYONPAS 132
Db 164 VNPGLIKSNDGPAA 178

RESULT 12
US-08-461-939B-72
Sequence 72, Application US/08461939B
Patent No. 6335019

GENERAL INFORMATION:
APPLICANT: Rogers, Bruce
APPLICANT: Klapper, David G.
APPLICANT: Rafnar, Thorunn
APPLICANT: Kuo, Mei-chang
TITLE OF INVENTION: Methods For Treating Sensitivity To A
TITLE OF INVENTION: Protein Allergen Using Peptides Which Include A T Cell Epitope
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,939B
FILING DATE:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/464,000
FILING DATE: 05-JUN-1995
APPLICATION NUMBER: US 08/290,448
FILING DATE: 15-AUG-1994
APPLICATION NUMBER: US 07/529,951
FILING DATE: 29-MAY-1990
APPLICATION NUMBER: US 07/325,365
FILING DATE: 17-MAR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-018CNDV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400

TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-939B-72

Query Match 13.9%; Score 98.5; DB 4; Length 387;
Best Local Similarity 34.7%; Pred. No. 0.00012;
Matches 26; Conservative 12; Mismatches 22; Indels 15; Gaps 3;

Qy 73 RPLWIFSGNNMIKIKMPYIAGYKTFDGRRAEVSYYH---VNGAK--FIRRD----- 121
Db 104 RPLWIFSGNNMIKIKMPYIAGYKTFDGRRAEVSYYH---VNGAK--FIRRD----- 163
Qy 122 ---GIIAAYONPAS 132
Db 164 VNPGLIKSNDGPAA 178

RESULT 13
US-08-464-000-72
Sequence 72, Application US/08464000
Patent No. 6335020

GENERAL INFORMATION:
APPLICANT: Rogers, Bruce
APPLICANT: Klapper, David G.
APPLICANT: Rafnar, Thorunn
APPLICANT: Kuo, Mei-chang
TITLE OF INVENTION: Allergenic Peptides from Ragweed Pollen
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,000
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/290,448
FILING DATE: 15-AUG-1994
APPLICATION NUMBER: US 07/529,951
FILING DATE: 29-MAY-1990
APPLICATION NUMBER: US 07/325,365
FILING DATE: 17-MAR-1989

ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-018CND2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-464-000-72

Query Match 13.9%; Score 98.5; DB 4; Length 387;
Best Local Similarity 34.7%; Pred. No. 0.00012;
Matches 26; Conservative 12; Mismatches 22; Indels 15; Gaps 3;

QY 73 RPLWIIISGNNNIIKMPYIAGYKTFDGRRAEVSYYH-----VNGAK--FIRVD----- 121
 DB 104 RPLWIIIFERDVIIRLDKEMVYVNSDKTIDGAGAVEIINAGFTLNGKNIYIHNINMHQDYK 163
 QY 122 -----GIIAAYONPAS 132
 DB 164 VNBGLIKSNDGPAA 178

RESULT 14

US-08-290-448A-76
 ; Sequence 76, Application US/08290448A
 ; Patent No. 5676954

GENERAL INFORMATION:

APPLICANT: Rogers, Bruce
 APPLICANT: Klapper, David G.
 APPLICANT: Ratnat, Thorunn
 APPLICANT: Kuo, Mei-chang
 TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses
 NUMBER OF SEQUENCES: 93
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LAHIVE & COCKFIELD
 STREET: 60 State Street, suite 510
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/290,448A
 FILING DATE: August 15, 1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/529,951
 FILING DATE: May 29, 1990
 APPLICATION NUMBER: US 07/325,365
 FILING DATE: March 17, 1989
 ATTORNEY/AGENT INFORMATION:
 NAME: Amy E. Mandragouras
 REGISTRATION NUMBER: 36,207
 REFERENCE/DOCKET NUMBER: IMI-018CN
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)227-7400
 TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 76:

SEQUENCE CHARACTERISTICS:
 LENGTH: 397 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-290-448A-76

Query Match 13.6%; Score 97; DB 1; Length 397;
 Best Local Similarity 29.5%; Pred. No. 0.0002;

Matches 31; Conservative 7; Mismatches 33; Indels 34; Gaps 3;

QY 52 WKNNRIWLO-----FAKLT-----GFTLMGRRPLWIIIF 79
 DB 61 WENNROALADCAQGFRAKTYGGKMGDVYTVTSNDDVANPEKGTLPFAAQNRPFWIIIF 120
 QY 80 SGNNNIKLMKMPYIAGYKTFDGRRAEVSYYHNGAKFIRVDGII 124
 DB 121 KNDWVININQELVNSDKTIDRGVKEI--INGGLTLMVKNII 163

RESULT 15

US-08-290-448A-76
 ; Sequence 76, Application US/08290448A
 ; Patent No. 5698204

GENERAL INFORMATION:
 APPLICANT: Rogers, Bruce
 APPLICANT: Klapper, David G.
 APPLICANT: Ratnat, Thorunn
 APPLICANT: Kuo, Mei-chang
 TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses
 NUMBER OF SEQUENCES: 93
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LAHIVE & COCKFIELD
 STREET: 60 State Street, suite 510
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/290,448A
 FILING DATE: August 15, 1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/529,951
 FILING DATE: May 29, 1990
 APPLICATION NUMBER: US 07/325,365
 FILING DATE: March 17, 1989
 ATTORNEY/AGENT INFORMATION:
 NAME: Amy E. Mandragouras
 REGISTRATION NUMBER: 36,207
 REFERENCE/DOCKET NUMBER: IMI-018CN
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)227-7400
 TELEFAX: (617)227-5941

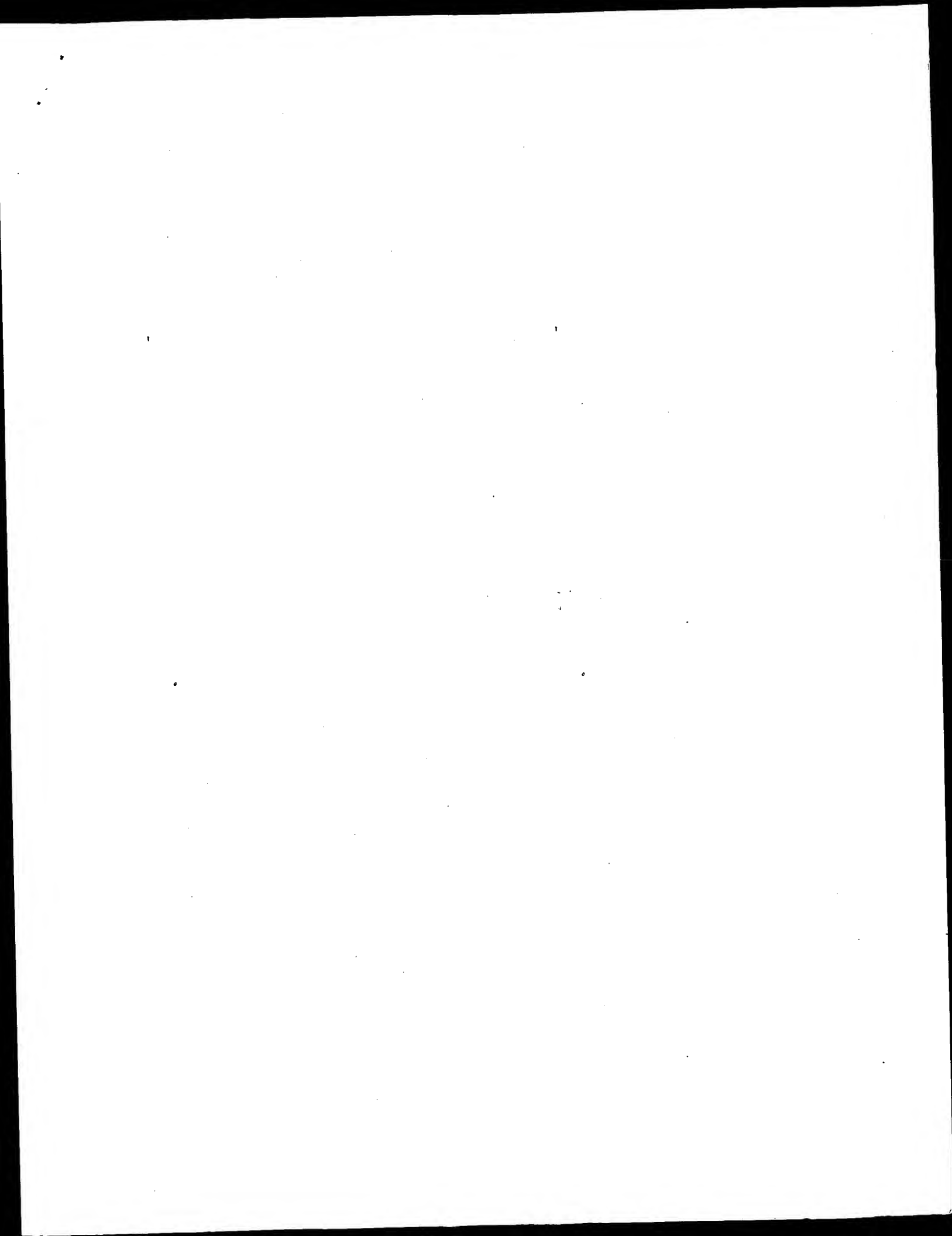
INFORMATION FOR SEQ ID NO: 76:

SEQUENCE CHARACTERISTICS:
 LENGTH: 397 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-290-448A-76

Query Match 13.6%; Score 97; DB 1; Length 397;
 Best Local Similarity 29.5%; Pred. No. 0.0002;
 Matches 31; Conservative 7; Mismatches 33; Indels 34; Gaps 3;

QY 52 WKNNRIWLO-----FAKLT-----GFTLMGRRPLWIIIF 79
 DB 61 WENNROALADCAQGFRAKTYGGKMGDVYTVTSNDDVANPEKGTLPFAAQNRPFWIIIF 120
 QY 80 SGNNNIKLMKMPYIAGYKTFDGRRAEVSYYHNGAKFIRVDGII 124
 DB 121 KNDWVININQELVNSDKTIDRGVKEI--INGGLTLMVKNII 163

Search completed: April 23, 2003, 18:05:34
 Job time : 17 secs



GenCore version 5.1.4 ps 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:56:55 / Search time 67.5789 Seconds
(without alignments)
243.919 Million cell updates/sec

Title: US-09-142-524D-1

Perfect score: 406
Sequence: 1 MKVTVAFNQFGPNRRVFIK.....IASRRVDGIIAAYQNPASWK 80

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	18.0	367	10 Q9L1T2	Q9L1T2 juniperus v
2	73	18.0	367	10 Q9L1T1	Q9L1T1 juniperus v
3	73	18.0	367	10 Q9M4S6	Q9M4S6 cupressus s
4	73	18.0	367	10 Q9M4S5	Q9M4S5 cupressus s
5	73	18.0	367	10 Q9M4S4	Q9M4S4 cupressus s
6	73	18.0	367	10 Q9M4S3	Q9M4S3 cupressus s
7	73	18.0	367	10 Q9M4S2	Q9M4S2 cupressus s
8	73	18.0	367	10 Q9M4S1	Q9M4S1 juniperus o
9	73	18.0	367	10 Q93XL6	Q93XL6 cupressus a
10	71	17.5	374	10 Q8RUB1	Q8RUB1 juniperus a
11	71	17.5	507	10 Q9FY19	Q9FY19 cryptomeria
12	70	17.2	152	13 Q90T05	Q90T05 juniperus a
13	69.5	17.1	335	2 Q9AHF2	Q9AHF2 ictalurus p
14	68	16.7	394	10 Q654S6	Q654S6 arabidopsis
15	67	16.5	674	16 Q9KDQ3	Q9KDQ3 bacillus ha
16	66.5	16.4	155	2 Q93SS5	Q93SS5 mycobacteri

17	66	16.3	350	16 P73865	P73865 synchocyst
18	66	16.3	394	10 Q654S7	Q654S7 arabidopsis
19	66	16.3	464	16 Q99W16	Q99W16 staphylococ
20	65.5	16.1	391	17 Q9YB91	Q9YB91 aeropyrum p
21	65.5	16.1	531	5 Q9TX24	Q9TX24 caenorhabdi
22	65.5	16.1	1758	16 Q9QMS5	Q9QMS5 escherichia
23	65	16.0	114	2 Q9X790	Q9X790 mycobacteri
24	65	16.0	132	5 Q962R1	Q962R1 spidoptera
25	65	16.0	390	10 Q65388	Q65388 arabidopsis
26	65	16.0	488	12 Q88525	Q88525 turkey herp
27	64.5	15.9	770	16 Q8YL18	Q8YL18 anabana sp
28	64	15.8	222	8 Q21280	Q21280 reclinomona
29	64	15.8	398	16 Q9WZL9	Q9WZL9 thermotoga
30	64	15.8	433	16 Q9K093	Q9K093 vibrio chol
31	63.5	15.6	533	5 Q01617	Q01617 caenorhabdi
32	63	15.5	645	16 Q98504	Q98504 rhizobium 1
33	63	15.5	753	13 Q9DE07	Q9DE07 gallus gall
34	62.5	15.4	629	4 Q9UPY1	Q9UPY1 homo sapien
35	62	15.3	244	16 Q9A100	Q9A100 streptococ
36	62	15.3	368	17 Q97Z14	Q97Z14 sulfobolus
37	62	15.3	559	4 Q9H796	Q9H796 homo sapien
38	62	15.3	567	16 Q91120	Q91120 pseudomonas
39	61.5	15.1	102	2 Q66004	Q66004 mycoplasma
40	61.5	15.1	228	17 Q977C1	Q977C1 sulfobolus
41	61.5	15.1	369	16 Q9RZU0	Q9RZU0 deinococcus
42	61.5	15.1	543	10 Q9PMV3	Q9PMV3 arabidopsis
43	61.5	15.1	730	2 Q9X571	Q9X571 rhizobium 1
44	61	15.0	142	5 Q9GPA9	Q9GPA9 chetax deat
45	61	15.0	173	2 Q47446	Q47446 escherichia

ALIGNMENTS

RESULT 1
ID Q9L1T2 PRELIMINARY; PRT; 367 AA.

AC 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Pollen major allergen 1-2.
OS Juniperus virginiana (Eastern red cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferales; Cupressaceae; Juniperus.
OX NCBI_TaxId=39584;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21315424; PubMed=11422137;
RA "Midoro-Horuti T.M., Goldblum R.M., Brooks E.G.;
RT "Identification of mutations in the genes for the pollen allergens of
RT eastern red cedar (Juniperus virginiana).";
RL Clin. Exp. Allergy 31:771-778 (2001).
DR EMBL; AF151427; AAFA80164.1; -
DR InterPro; IPR002022; Amb allergen.
DR Pfam; PF00544; pec lyase; 1.
DR PRINTS; PRO0807; AMBAL1ERGN.
SQ SEQUENCE 367 AA; 39768 MW; 0A6AC2F1BAF89586 CRC64;

Query Match 18.0%; Score 73; DB 10; Length 367;
Best Local Similarity 55.6%; Pred. No. 2.3;
Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKVTVAFNQFGPNRRVFIKRVSVITH 27
DB 234 MKVTVAFNQFGPNAGGMPRRARYGLVH 260

RESULT 2
ID Q9L1T1 PRELIMINARY; PRT; 367 AA.
AC 01-OCT-2000 (T-EMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Pollen major allergen 1-1.
 OS Juniperus virginiana (Eastern red cedar).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
 OK NCBI_TaxID=39584;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21315424; PubMed=11422137;
 RA Midoro-Horiuchi T.M., Goldblum R.M., Brooks E.G.;
 RT "Identification of mutations in the genes for the pollen allergens of
 eastern red cedar (Juniperus virginiana).";
 RL Clin. Exp. Allergy 31:771-778(2001).
 DR EMBL; AF151429; AAF80166.1; -
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec_lyase1.1.
 DR PRINTS; PR00807; AMBAALLERGEN.
 SQ SEQUENCE 367 AA; 39708 MW; DCBD1981A74E711 CRC64;

Query Match 18.0%; Score 73; DB 10; Length 367;
 Best Local Similarity 55.6%; Pred. No. 2.3;
 Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKYTVAFNQGPNRRVFIKRVSNVITH 27
 DB 234 MKYTVAFNQGPNAGQMPRPRARYGLVH 260

RESULT 3

ID Q9M4S6 PRELIMINARY; PRT; 367 AA.

DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUP81.
 OS Cupressus sempervirens.
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OK NCBI_TaxID=13469;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Monsalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens.";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257491; AAF72625.1; -
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec_lyase1.1.
 DR PRINTS; PR00807; AMBAALLERGEN.
 SQ SEQUENCE 367 AA; 39835 MW; B4E9C60108C2C5A3 CRC64;

Query Match 18.0%; Score 73; DB 10; Length 367;
 Best Local Similarity 55.6%; Pred. No. 2.3;
 Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKYTVAFNQGPNRRVFIKRVSNVITH 27
 DB 234 MKYTVAFNQGPNAGQMPRPRARYGLVH 260

RESULT 4
 ID Q9M4S5 PRELIMINARY; PRT; 367 AA.
 AC Q9M4S5;
 DT 01-OCT-2000 (T-EMBlrel. 15, Created)
 DT 01-OCT-2000 (T-EMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (T-EMBlrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUP81.
 OS Cupressus sempervirens.

CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OK NCBI_TaxID=13469;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Monsalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens.";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257492; AAF72626.1; -
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec_lyase1.1.
 DR PRINTS; PR00807; AMBAALLERGEN.
 SQ SEQUENCE 367 AA; 39894 MW; 5D56FC0E3263B741 CRC64;

Query Match 18.0%; Score 73; DB 10; Length 367;
 Best Local Similarity 55.6%; Pred. No. 2.3;
 Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKYTVAFNQGPNRRVFIKRVSNVITH 27
 DB 234 MKYTVAFNQGPNAGQMPRPRARYGLVH 260

RESULT 5

ID Q9M4S4 PRELIMINARY; PRT; 367 AA.

DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUP81.
 OS Cupressus sempervirens.
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OK NCBI_TaxID=13469;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Monsalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens.";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257493; AAF72627.1; -
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec_lyase1.1.
 DR PRINTS; PR00807; AMBAALLERGEN.
 SQ SEQUENCE 367 AA; 39934 MW; 974D3011D74E3D6E CRC64;

Query Match 18.0%; Score 73; DB 10; Length 367;
 Best Local Similarity 55.6%; Pred. No. 2.3;
 Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKYTVAFNQGPNRRVFIKRVSNVITH 27
 DB 234 MKYTVAFNQGPNAGQMPRPRARYGLVH 260

RESULT 6
 ID Q9M4S3 PRELIMINARY; PRT; 367 AA.
 AC Q9M4S3;
 DT 01-OCT-2000 (T-EMBlrel. 15, Created)
 DT 01-OCT-2000 (T-EMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (T-EMBlrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUP81.
 OS Cupressus sempervirens.
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OK NCBI_TaxID=13469;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Monsalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens.";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF257494; AAF72628.1; -
 DR InterPro: IPR002022; Amb_allergen.
 DR Pfam: PF00544; pec_lyase.1.
 DR PRINTS: PR00807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39832 MW; B5DFB5A61C07A53 CRC64;

Query Match 18.0%; Score 73; DB 10; Length 367;
 Best Local Similarity 55.6%; Pred. No. 2.3;
 Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKVTAFNQGPNRRVFIKRVSNVITH 27
 DB 234 MKVTAFNQGPNRRVFIKRVSNVITH 260

RESULT 7
 ID Q9M4S2 PRELIMINARY; PRT; 367 AA.
 AC Q9M4S2;
 DT 01-OCT-2000 (T-EMBLrel. 15, Created)
 DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUPSI.
 OS Cupressus sempervirens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OX NCBI_TaxID=13469;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Monsalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens.";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF257495; AAF72629.1; -
 DR InterPro: IPR002022; Amb_allergen.
 DR Pfam: PF00544; pec_lyase.1.
 DR PRINTS: PR00807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39819 MW; AF7E055A61C07A53 CRC64;

Query Match 18.0%; Score 73; DB 10; Length 367;
 Best Local Similarity 55.6%; Pred. No. 2.3;
 Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKVTAFNQGPNRRVFIKRVSNVITH 27
 DB 234 MKVTAFNQGPNRRVFIKRVSNVITH 260

RESULT 8
 ID Q93X51 PRELIMINARY; PRT; 367 AA.
 AC Q93X51;
 DT 01-DEC-2001 (T-EMBLrel. 19, Created)
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
 DE Putative allergen jun o 1.
 GN JUN O 1.
 OS Juniperus oxycedrus (Prickly juniper).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
 OX NCBI_TaxID=69008;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE=POLLEN;
 RA Iacovacci P., Di Felice G., Pini C.;
 RT "Cloning of Juniperus oxycedrus major allergen.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ293767; CAC48400.1; -

DR InterPro: IPR002022; Amb_allergen.
 DR Pfam: PF00544; pec_lyase.1.
 SQ SEQUENCE 367 AA; 39808 MW; 5D28204DBFDB9D7 CRC64;

Query Match 18.0%; Score 73; DB 10; Length 367;
 Best Local Similarity 55.6%; Pred. No. 2.3;
 Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKVTAFNQGPNRRVFIKRVSNVITH 27
 DB 234 MKVTAFNQGPNRRVFIKRVSNVITH 260

RESULT 9
 ID Q93XL6 PRELIMINARY; PRT; 367 AA.
 AC Q93XL6;
 DT 01-DEC-2001 (T-EMBLrel. 19, Created)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
 DE Putative allergen Cup a 1 precursor.
 GN CUP A 1.
 OS Cupressus arizonica.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OX NCBI_TaxID=49011;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE=POLLEN;
 RA Buteront C., Di Felice G., Pini C.;
 RT "Cloning of Cupressus arizonica major allergen.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ278498; CAC37790.2; -
 DR Signal.
 KM SIGNAL. 1 21
 FT CHAIN 22 367 POTENTIAL.
 SQ SEQUENCE 367 AA; 39809 MW; AFE97260423A928 CRC64;

Query Match 18.0%; Score 73; DB 10; Length 367;
 Best Local Similarity 55.6%; Pred. No. 2.3;
 Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKVTAFNQGPNRRVFIKRVSNVITH 27
 DB 234 MKVTAFNQGPNRRVFIKRVSNVITH 260

RESULT 10
 ID Q8RUR1 PRELIMINARY; PRT; 374 AA.
 AC Q8RUR1;
 DT 01-JUN-2002 (T-EMBLrel. 21, Created)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
 DE CRY J 1 precursor.
 GN CRY J 1.1 OR CRY J 1.2.
 OS Cryptomeria japonica (Japanese cedar).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.
 OX NCBI_TaxID=3369;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE=POLLEN;
 RA Futamura N., Shinohara K.;
 RT "Isolation and characterization of cDNAs encoding major allergen Cry j
 1 from Cryptomeria japonica pollen.";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB081309; BAB86286.1; -
 DR EMBL: AB081310; BAB86287.1; -
 KW Signal.
 FT CHAIN 1 21 POTENTIAL.
 FT SIGNAL 22 374 CRY J 1.
 SQ SEQUENCE 374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;

Query Match 17.5%; Score 71; DB 10; Length 374;
 Best Local Similarity 55.6%; Pred. No. 4.1;
 Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKVTAFNPGFNRVRVFKRVSNVTH 27
 |||||
 DB 234 MKVTAFNPGFNRVRVFKRVSNVTH 260

RESULT 11

Q9FY19 PRELIMINARY; PRT; 507 AA.

AC Q9FY19
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Pollen major allergen 2 protein precursor.
 GN JNA2.
 OS Juniperus ashei (Ozark white cedar).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
 OK NCBI_TaxID=13101;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MALE POLLEN;
 RX MEDLINE=20403896; PubMed=10944464;
 RA Yochiyama M., Miyahara M., Shimizu K., Kino K., Tsunoo H.;
 RT "Purification, identification and cDNA cloning of Jun a 2, the second
 major allergen of mountain cedar pollen."
 RL Biochem. Biophys. Res. Commun. 275:195-202(2000).
 CC -1 SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
 CC (POLYGALACTURONASES).
 CC EMBL: AJ404653; CAC05582.1; -
 DR HSPB; P26509; 188.
 DR InterPro; IPR000743; GH28.
 DR InterPro; IPR000408; Reg_chir_condens.
 DR Pfam; PF00295; Glyco_hydro_28; 1.
 DR PROSITE; PS00502; POLYGALACTURONASE; UNKNOWN_1.
 DR PROSITE; PS00626; RCCL_2; UNKNOWN_1.
 KW Cell wall; Glycosidase; Hydrolase; Signal.
 FT SIGNAL 1 54
 FT POTENTIAL.
 SQ SEQUENCE 507 AA; 55730 MW; 28280AA5E358FE5A CRC64;

Query Match 17.5%; Score 71; DB 10; Length 507;
 Best Local Similarity 73.7%; Pred. No. 5.9;
 Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 31 IDIFASKNFHLQKNTIGTG 49
 |||||
 DB 237 IDIFASKNFHLQKNTIGTG 255

RESULT 12

Q90Y05 PRELIMINARY; PRT; 152 AA.

AC Q90Y05
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE 40S ribosomal protein S18.
 OS Ictalurus punctatus (Channel catfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
 OC Ictaluridae; Ictalurus.
 OK NCBI_TaxID=7998;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Karsl A., Patterson A., Feng J., Liu Z.J.;
 RT "Translational machinery of channel catfish: I. A transcriptomic
 approach to the analysis of 32 40S ribosomal protein genes and their
 expression."
 RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF402827; AAK95201.1; -
 DR InterPro; IPR001892; Ribosomal_S13.
 DR Pfam; PF00416; Ribosomal_S13; 1.
 DR ProDom; PD001363; Ribosomal_S13; 1.
 DR PROSITE; PS00646; RIBOSOMAL_S13; UNKNOWN_1.
 KW Ribosomal protein.
 SQ SEQUENCE 152 AA; 17691 MW; 1F018266809D92A1 CRC64;

Query Match 17.2%; Score 70; DB 13; Length 152;
 Best Local Similarity 29.1%; Pred. No. 1.9;
 Matches 23; Conservative 8; Mismatches 32; Indels 16; Gaps 2;

QY 12 PRRRVFTRVGNVTHGRIDIFASKNFHLQKNTIGTGR-----ISLKTSGKI 61
 |||||
 DB 6 PRRFHLRLVNTNIDGKRKTAFAITAK-----GVGRRAHVVLKADIDLSKRGEL 59

QY 62 ASRRVDGIITAAVONPASKW 80
 |||||
 DB 60 TDEVERVVTIMQNPQYK 78

RESULT 13

Q9AHP2 PRELIMINARY; PRT; 335 AA.

AC Q9AHP2
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Hypothetical 36.5 kDa protein.
 OS Agrobacterium tumefaciens.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OK NCBI_TaxID=358;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D3;
 RX MEDLINE=21322708; PubMed=11429459;
 RA Trost S., Bauer R., Knackmuss H.J., Stolz A.;
 RT "Genetic and biochemical characterization of an enantioselective
 amylase from Agrobacterium tumefaciens strain d3."
 RL Microbiology 147:1815-1824(2001).
 RL EMBL; AF315580; AAK28494.1; -
 DR InterPro; IPR000620; DUF6.
 DR Pfam; PF00892; DUF6; 2.
 KW Hypothetical protein.
 SQ SEQUENCE 335 AA; 36505 MW; 334C9B65A972C0C1 CRC64;

Query Match 17.1%; Score 69.5; DB 2; Length 335;
 Best Local Similarity 38.6%; Pred. No. 5.4;
 Matches 27; Conservative 4; Mismatches 28; Indels 11; Gaps 3;

QY 10 FGRNRVFRVGNVTHGRIDIFASKNFHLQKNTIGT-----GRISLKTSGKIA 62
 |||||
 DB 240 FGRVATSFCEVVAEAV--GRRISVFAMSNLTGVPITGAIAIFGRKLSVLSGFWL 297

QY 63 SRRVDGIITAA 72
 |||||
 DB 298 I--VSGIATA 305

RESULT 14

Q65456 PRELIMINARY; PRT; 394 AA.

AC Q65456
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Pectate lyase like protein.
 GN FIN20.180 OR AT4G22080.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Bevan M., Wedler H., Wamburt R., Bancroft I., Mewes H.W., Mayer K.,
 RA Scheller C.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Wedler H., Wamburt R., Mewes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL022140; CA1811.1; -;
 DR EMBL; AL161556; CAB79163.1; -;
 DR InterPro; IPR002022; Amb allergen.
 DR InterPro; IPR002048; EF-Hand.
 DR Pfam; PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALERGEN.
 DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.
 KW Lyase.
 SQ SEQUENCE 394 AA; 43476 MW; A48D586ECF148CA CRC64;

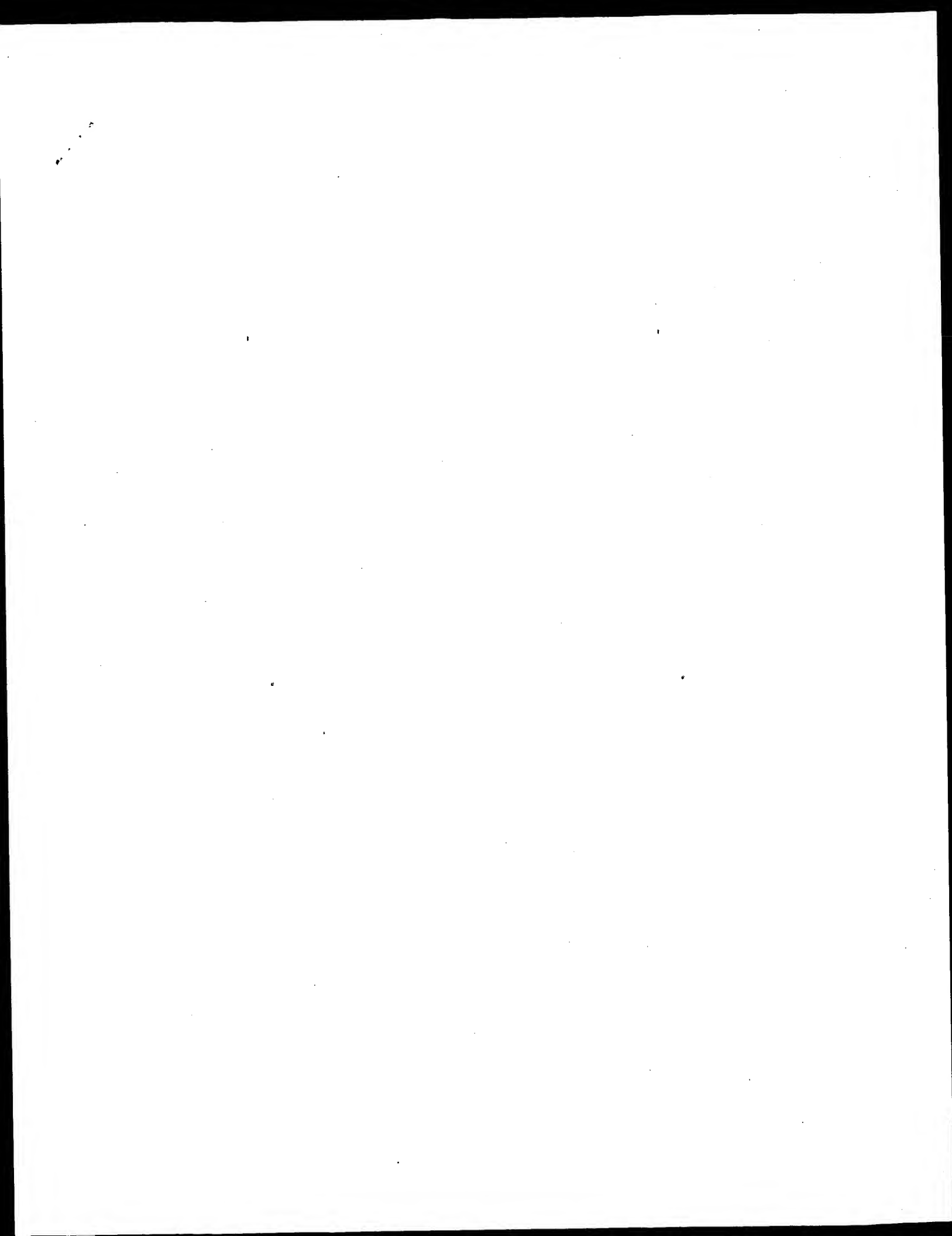
Query Match 16.7%; Score 68; DB 10; Length 394;
 Best Local Similarity 28.6%; Pred. No. 9.8;
 Matches 24; Conservative 7; Mismatches 27; Indels 26; Gaps 3;
 QY 1 MKVTAFNPGPNRRVFIKVSNTI--GRID-----IFASKNPHLOKN 44
 DB 256 MKVTAFNPGPNRRVFIKVSNTI--GRID-----IFASKNPHLOKN 44
 QY 45 TIGTRRISLKTSGKIASRRVDG 68
 DB 316 -----KSNKSVTKREVKG 329

RESULT 15
 ID 09KD03 PRELIMINARY; PRT; 674 AA.
 AC 09KD03;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE ABC transporter (ATP-binding protein).
 GN BH1158.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Bacillaceae; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Maui N.,
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kohara S.,
 RA Horiuchi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL; AP001511; BAB04877.1; -;
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR001140; ABCTransporterTM.
 DR InterPro; IPR003439; ABC transporter.
 DR Pfam; PF00664; ABC membrane; 1.
 DR Pfam; PF00005; ABC_tran; 1.
 DR ProDom; PD000006; ABC_transport; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_1.
 KW ATP-binding; Complete proteome.
 SQ SEQUENCE 674 AA; 75735 MW; 6AF94B058C84C957 CRC64;

Best Local Similarity 29.7%; Pred. No. 24;
 Matches 19; Conservative 15; Mismatches 20; Indels 10; Gaps 3;
 QY 1 MKVTAFNPGPNRRVFIKVSNTI--GRIDIFASKNPHLOKNITIGTRRISLKTSGK 60
 DB 158 LMVVASFFQYQ--ORFYLQKAVNRITQRLIDLFN-----HLSRLPV---RFDMMPACK 207
 QY 61 IASR 64
 DB 208 VVSR 211

Search completed: April 20, 2003, 13:12:54
 Job time : 71.5789 secs

Query Match 16.5%; Score 67; DB 16; Length 674;



GenCore version 5.1.4.P5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:05 ; Search time 97.0526 Seconds
(without alignments)
109,838 Million cell updates/sec

Title: US-09-142-524D-1

Perfect score: 406

Sequence: 1 MKYVAFNQFGNRRVFIK.....IASRVYDGIIAQNPASWK 80

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /SID2/gcgdata/geneeq/geneeq-emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneeq/geneeq-emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneeq/geneeq-emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneeq/geneeq-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneeq/geneeq-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneeq/geneeq-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneeq/geneeq-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneeq/geneeq-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneeq/geneeq-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneeq/geneeq-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneeq/geneeq-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneeq/geneeq-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneeq/geneeq-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneeq/geneeq-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneeq/geneeq-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneeq/geneeq-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneeq/geneeq-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneeq/geneeq-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneeq/geneeq-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneeq/geneeq-emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneeq/geneeq-emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneeq/geneeq-emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneeq/geneeq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	406	100.0	80	18	AAW27369
2	327.5	80.7	105	18	AAW27370
3	320	78.8	134	18	AAW27371
4	174.5	43.0	214	22	AAW69120
5	172	42.4	186	22	AAW69102
6	164	40.4	210	22	AAW68103
7	144	35.5	47	19	AAW80353
8	140.5	34.6	74	21	AAW23905
9	140.5	34.6	94	21	AAW23906
10	140.5	34.6	95	21	AAW23897

11	140.5	34.6	95	21	AAW23901	Artificial sequenc
12	140.5	34.6	95	22	AAW69098	Cedar pollen aller
13	140.5	34.6	96	22	AAW69104	Cedar pollen aller
14	140.5	34.6	96	22	AAW69119	Cedar pollen aller
15	140.5	34.6	97	22	AAW69105	Cedar pollen aller
16	140.5	34.6	99	21	AAW23902	Artificial sequenc
17	140.5	34.6	99	21	AAW23907	Artificial sequenc
18	140.5	34.6	99	21	AAW69099	Cedar pollen aller
19	135	33.3	81	19	AAW80358	Sugi allergen prot
20	135	33.3	81	21	AAW23895	Artificial sequenc
21	135	33.3	81	22	AAW69118	Cedar pollen aller
22	135	33.3	88	22	AAW69114	Cedar pollen aller
23	135	33.3	93	21	AAW23881	Artificial sequenc
24	135	33.3	93	21	AAW23896	Artificial sequenc
25	135	33.3	93	22	AAW69097	Cedar pollen aller
26	135	33.3	93	22	AAW69117	Cedar pollen aller
27	133	32.8	71	21	AAW23878	Artificial sequenc
28	133	32.8	81	21	AAW23874	Artificial sequenc
29	133	32.8	81	21	AAW23879	Artificial sequenc
30	133	32.8	81	22	AAW69094	Artificial sequenc
31	127.5	31.4	79	21	AAW23880	Cedar pollen aller
32	127.5	31.4	79	21	AAW23884	Artificial sequenc
33	127.5	31.4	79	22	AAW69095	Artificial sequenc
34	125.5	30.9	47	19	AAW80352	Cedar pollen aller
35	125.5	30.9	79	21	AAW23885	Sugi allergen prot
36	125.5	30.9	79	21	AAW23880	Artificial sequenc
37	125.5	30.9	79	22	AAW69096	Artificial sequenc
38	121.5	29.9	47	19	AAW80351	Cedar pollen aller
39	120	29.6	47	19	AAW80357	Sugi allergen prot
40	114	28.1	47	19	AAW80356	Immunomodulatory p
41	103	25.4	42	21	AAW84105	Artificial sequenc
42	102	25.1	61	21	AAW23894	Artificial sequenc
43	100.5	24.8	61	21	AAW23883	Artificial sequenc
44	99	24.4	19	22	AAW84115	Amino acid sequenc
45	99	24.4	16	22	AAW69791	Japanicum allergen

ALIGNMENTS

RESULT 1
AAW27369 standard; peptide; 80 AA.
XX
AC AAW27369;
XX
DT 24-MAR-1998 (first entry)
XX
DB Multi-epitope peptide used as immunotherapeutic agent #1.
XX
KM Multi-epitope peptide; immunotherapeutic agent; allergic disease;
KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
XX
OS Synthetic.
XX
PN WO9732600-A1.
XX
PD 12-SEP-1997.
XX
PF 10-MAR-1997; 97WO-UP00740.
XX
PR 10-MAR-1996; 96JP-0080702.
XX
PA (MEIP) MELI MILK PROD CO LTD.
XX
PI Daiiriki K, Iwama A, Kino K, Kume A, Sone T;
XX
XX WPI; 1997-470495/43.
XX
PT Peptide immuno-therapeutic agent to treat allergic diseases -
PT contains multi-epitope peptide containing T cell epitope regions
PT from different allergens
XX

PS Claim 6; Page 31; 58pp; Japanese.

XX The present sequence represents a multi-epitope peptide which is used as
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 CC or more different allergens (preferably linked via arginine or lysine
 CC dimers), where the T cell epitope regions have a positivity index
 CC greater than 100 as measured in a patient group responding to the
 CC allergen; have at least 70% reactivity with lymphocytes from patients
 CC responding to the allergen; and are not reactive with immunoglobulin E
 CC (IgE) antibodies from patients responsive to the allergen. The agent can
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.

SQ Sequence 80 AA;

Query Match 100.0%; Score 406; DB 18; Length 80;
 Best Local Similarity 100.0%; Pred. No. 2,6e-48;
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVTVAFNQFGPNRRVFYIKRVSNVTHGRRIDIFASKNFHLQKNTIGGRISLKLTSCK 60
 DB 1 MKVTVAFNQFGPNRRVFYIKRVSNVTHGRRIDIFASKNFHLQKNTIGGRISLKLTSCK 60
 QY 61 IASRRVDGIIAAYQNPASWK 80
 DB 61 IASRRVDGIIAAYQNPASWK 80

RESULT 2
 AAM27370
 ID AAM27370 standard; peptide; 105 AA.
 XX AAM27370;
 AC AAM27370;
 XX 24-MAR-1998 (first entry)
 DT 24-MAR-1998 (first entry)
 XX Multi-epitope peptide used as immunotherapeutic agent #2.
 DE Multi-epitope peptide; immunotherapeutic agent; allergic disease;
 XX T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
 KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;
 OS Synthetic.
 XX WO9732600-A1.
 PN WO9732600-A1.
 XX 12-SEP-1997.
 PD 12-SEP-1997.
 XX 10-MAR-1997; 97WO-JP00740.
 PF 10-MAR-1997; 97WO-JP00740.
 XX 10-MAR-1996; 96JP-0080702.
 PR 10-MAR-1996; 96JP-0080702.
 XX (MEIP) MEIJI MILK PROD CO LTD.
 PA Dairiki K, Iwama A, Kuno K, Kume A, Sone T;
 PI Dairiki K, Iwama A, Kuno K, Kume A, Sone T;
 XX WPI; 1997-470495/43.
 DR WPI; 1997-470495/43.
 XX Peptide immuno-therapeutic agent to treat allergic diseases -
 PT contains multi-epitope peptide containing T cell epitope regions
 PT from different allergens
 XX
 PS Claim 6; Page 31; 58pp; Japanese.

XX The present sequence represents a multi-epitope peptide which is used as
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 CC or more different allergens (preferably linked via arginine or lysine
 CC dimers), where the T cell epitope regions have a positivity index
 CC greater than 100 as measured in a patient group responding to the
 CC allergen; have at least 70% reactivity with lymphocytes from patients
 CC responding to the allergen; and are not reactive with immunoglobulin E
 CC (IgE) antibodies from patients responsive to the allergen. The agent can
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.

XX SQ Sequence 105 AA;

Query Match 80.7%; Score 327.5; DB 18; Length 105;
 Best Local Similarity 67.6%; Pred. No. 2,7e-37;
 Matches 71; Conservative 1; Mismatches 8; Indels 25; Gaps 2;

QY 1 MKVTVAFNQFGPNRRVFYIKRVSNVTHGRRIDIFASKNFHLQKNTIGGRISLKLTSCK 53
 DB 1 MKVTVAFNQFGPNRRVFYIKRVSNVTHGRRIDIFASKNFHLQKNTIGGRISLKLTSCK 53
 QY 54 -LKLTSCKIA-----SRVDGIIAAYQNPASWK 80
 DB 61 PAKLTGFTLMGRRLKMPMTIAGKTFDGRRDGIIAAYQNPASWK 105

RESULT 3
 AAM27371
 ID AAM27371 standard; peptide; 134 AA.
 XX AAM27371;
 AC AAM27371;
 XX 24-MAR-1998 (first entry)
 DT 24-MAR-1998 (first entry)
 XX Multi-epitope peptide used as immunotherapeutic agent #3.
 DE Multi-epitope peptide; immunotherapeutic agent; allergic disease;
 XX T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
 KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;
 OS Synthetic.
 XX WO9732600-A1.
 PN WO9732600-A1.
 XX 12-SEP-1997.
 PD 12-SEP-1997.
 XX 10-MAR-1997; 97WO-JP00740.
 PF 10-MAR-1997; 97WO-JP00740.
 XX 10-MAR-1996; 96JP-0080702.
 PR 10-MAR-1996; 96JP-0080702.
 XX (MEIP) MEIJI MILK PROD CO LTD.
 PA Dairiki K, Iwama A, Kuno K, Kume A, Sone T;
 PI Dairiki K, Iwama A, Kuno K, Kume A, Sone T;
 XX WPI; 1997-470495/43.
 DR WPI; 1997-470495/43.
 XX Peptide immuno-therapeutic agent to treat allergic diseases -
 PT contains multi-epitope peptide containing T cell epitope regions
 PT from different allergens
 XX
 PS Claim 6; Page 32; 58pp; Japanese.

XX The present sequence represents a multi-epitope peptide which is used as
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 CC or more different allergens (preferably linked via arginine or lysine
 CC dimers), where the T cell epitope regions have a positivity index
 CC greater than 100 as measured in a patient group responding to the
 CC allergen; have at least 70% reactivity with lymphocytes from patients
 CC responding to the allergen; and are not reactive with immunoglobulin E
 CC (IgE) antibodies from patients responsive to the allergen. The agent can
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.

SQ Sequence 134 AA;

Query Match 78.8%; Score 320; DB 18; Length 134;
 Best Local Similarity 53.7%; Pred. No. 4e-36;
 Matches 72; Conservative 2; Mismatches 6; Indels 54; Gaps 2;

QY 1 MKVTVAFNQFGPNRRVFYIKRVSNVTHGRRIDIFASKNFHLQKNTIGGRISLKLTSCK 51
 DB 1 MKVTVAFNQFGPNRRVFYIKRVSNVTHGRRIDIFASKNFHLQKNTIGGRISLKLTSCK 51
 QY 52 -----ISLKL-----TSGLIASRRV 66

Db 61 FAKTGFTLMGRRLPILIFSGNMNLIKMPYIAGYKTFDGRRAEVSXVHNGAKFIRRV 120
 QY 67 DGIIAAVONPASWK 80
 Db 121 DGIIAAVONPASWK 134

RESULT 4
 AAB69120
 ID AAB69120 standard; Protein; 214 AA.

AC AAB69120;

DT 23-APR-2001 (first entry)

DE Cedar pollen allergen T cell epitope derived protein SEQ ID NO:61.

KM Japanese cedar; Cryptomeria japonica; cedar pollen allergen;
 KM T cell epitope; anti-sugipollinosis.

OS Cryptomeria japonica.
 OS Synthetic.

PN JP2000327699-A.

PD 28-NOV-2000.

PF 15-MAR-2000; 2000JP-0071710.

PR 15-MAR-1999; 99JP-0068316.

PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA (SANY) SANKYO CO LTD.

DR WPI; 2001-185061/19.

DR N-PSDB; AAF59044.

PT Novel peptide and its use

PS Example 11; Page 58-59; 75pp; Japanese.

CC The present invention describes a peptide, its complex, derivative or
 CC its polymerizate, where the peptide (1) has a formula of:
 CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7
 CC where alpha1 to alpha7 = amino acid sequence selected from amino acid
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes
 CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen
 CC allergens. The peptide can be used in an anti-sugipollinosis agent.
 CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used
 CC in the exemplification of the present invention.

SQ Sequence 214 AA;

Query Match 43.0%; Score 174.5; DB 22; Length 214;
 Best Local Similarity 38.3%; Pred. No. 9.9e-16;

Matches 49; Conservative 0; Mismatches 2; Indels 77; Gaps 4;

QY 1 MKVTVAFNQFGPNRRVFIKRVSNVLIHGRRIDIFASKNFHLQKNTIGTGRRISLKTSGK 60
 Db 30 MKVTVAFNQFGP-----DIFASKNFHLQKN-----KLTSGK 60

QY 61 IAS-----RRVD--GIITAA 72
 Db 61 IASCLNYGLVHANNYDPSGKYEKGNITKKEAFNVEQFAKLTGFTLMGRADPRGIITAA 120

QY 73 YONPASWK 80
 Db 121 YONPASWK 128

RESULT 5
 AAB69102

ID AAB69102 standard; Protein; 186 AA.

AC AAB69102;

DT 23-APR-2001 (first entry)

DE Cedar pollen allergen T cell epitope derived protein SEQ ID NO:10.

KM Japanese cedar; Cryptomeria japonica; cedar pollen allergen;
 KM T cell epitope; anti-sugipollinosis.

OS Cryptomeria japonica.
 OS Synthetic.

PN JP2000327699-A.

PD 28-NOV-2000.

PF 15-MAR-2000; 2000JP-0071710.

PR 15-MAR-1999; 99JP-0068316.

PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA (SANY) SANKYO CO LTD.

DR WPI; 2001-185061/19.

DR N-PSDB; AAF59011.

PT Novel peptide and its use

PS Claim 14; Page 36-37; 75pp; Japanese.

CC The present invention describes a peptide, its complex, derivative or
 CC its polymerizate, where the peptide (1) has a formula of:
 CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7
 CC where alpha1 to alpha7 = amino acid sequence selected from amino acid
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes
 CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen
 CC allergens. The peptide can be used in an anti-sugipollinosis agent.
 CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used
 CC in the exemplification of the present invention.

SQ Sequence 186 AA;

Query Match 42.4%; Score 172; DB 22; Length 186;
 Best Local Similarity 40.9%; Pred. No. 1.8e-15;

Matches 47; Conservative 0; Mismatches 2; Indels 66; Gaps 4;

QY 1 MKVTVAFNQFGPNRRVFIKRVSNVLIHGRRIDIFASKNFHLQKNTIGTGRRISLKTSGK 60
 Db 29 MKVTVAFNQFGP-----FASKNFHLQKN-----KLTSGK 58

QY 61 IAS-----RRVD--GIITAAVONPASW 79
 Db 59 IASCLNYGLVHANNYDPSGKYEKGNITKKEAFNVEADPRGIITAAVONPASW 113

RESULT 6

AAB69103
 ID AAB69103 standard; Protein; 210 AA.

AC AAB69103;

DT 23-APR-2001 (first entry)

DE Cedar pollen allergen T cell epitope derived protein SEQ ID NO:12.

KM Japanese cedar; Cryptomeria japonica; cedar pollen allergen;
 KM T cell epitope; anti-sugipollinosis.

OS Cryptomeria japonica.
 OS Synthetic.

PN JP2000327699-A.
 XX
 PD 28-NOV-2000.
 XX
 PF 15-MAR-2000; 2000JP-00717710.
 XX
 PR 15-MAR-1999; 99JP-0068316.
 XX
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA (SANY) SANKYO CO LTD.
 XX
 DR WPI; 2001-185061/19.
 DR N-PSDB; AAF59012.
 XX
 PT Novel peptide and its use -
 XX
 PS Claim 7; Page 39-40; 75pp; Japanese.
 XX
 CC The present invention describes a peptide, its complex, derivative or
 CC its polymerizate, where the peptide (I) has a formula of:
 CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7
 CC where alpha1 to alpha7 = amino acid sequence selected from amino acid
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes
 CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen
 CC allergens. The peptide can be used in an anti-sugipollinosis agent.
 CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used
 CC in the exemplification of the present invention.
 XX
 SQ Sequence 210 AA;
 XX
 Query Match 40.4%; Score 164; DB 22; Length 210;
 Best Local Similarity 37.0%; Pred. No. 2.7e-14;
 Matches 47; Conservative 0; Mismatches 2; Indels 78; Gaps 4;
 XX
 QY 1 MKVTAFNFGPGRRRVFIRKSVNVIHGRIDIFASKNFHLQKNTIGTGRISLKTSGK 60
 DB 29 MKVTAFNFGP-----FASKNFHLQKNT-----KLTSGK 58
 QY 61 IAS-----RRVD-GITAA 72
 DB 59 IASCLNVLGVHANNVNDPSGKYEKGNITKKEAFNVEQFAKLTGFTLMGRADPRGIITAA 118
 QY 73 YONPASW 79
 DB 119 YONPASW 125
 XX
 RESULT 7
 AAM80353
 ID AAM80353 standard; peptide; 47 AA.
 XX
 AC AAM80353;
 XX
 DT 11-JAN-1999 (first entry)
 XX
 DE Sugi allergen protein Cryj1 derived epitope for T cells.
 XX
 KM T cell epitope; sugi allergen proteins Cryj1; Cryj2; treatment;
 KM sugi-pollinosis; allergic reaction; pollen.
 XX
 OS Synthetic.
 XX
 PN JP10259198-A.
 XX
 PD 29-SEP-1998.
 XX
 PF 22-DEC-1997; 97JP-0353448.
 XX
 PR 24-DEC-1996; 96JP-0343441.
 XX
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA (SANY) SANKYO CO LTD.
 XX

DR WPI; 1998-577037/49.
 XX
 PT A linked T cell epitope peptide - used for the treatment of
 PT sugi-pollinosis
 XX
 PS Claim 10; Page 5; 21pp; Japanese.
 XX
 CC AAM80339-58 represent epitopes for T cells, derived from the sugi
 CC allergen proteins Cryj1 (AAM80339-44, AAM80350-53 and AAM80356-58) and
 CC Cryj2 (AAM80345-49 and AAM80354-55). The peptides are useful for the
 CC treatment of sugi-pollinosis, an allergic reaction of the body to
 CC pollen.
 XX
 SQ Sequence 47 AA;
 XX
 Query Match 35.5%; Score 144; DB 19; Length 47;
 Best Local Similarity 44.3%; Pred. No. 2.3e-12;
 Matches 35; Conservative 0; Mismatches 0; Indels 44; Gaps 2;
 XX
 QY 1 MKVTAFNFGPGRRRVFIRKSVNVIHGRIDIFASKNFHLQKNTIGTGRISLKTSGK 60
 DB 13 MKVTAFNFGP-----FASKNFHLQKNT----- 35
 QY 61 IASRRVDGITAAVONPASW 79
 DB 36 -----GITAAVONPASW 47
 XX
 RESULT 8
 AAB23905
 ID AAB23905 standard; peptide; 74 AA.
 XX
 AC AAB23905;
 XX
 DT 17-JAN-2001 (first entry)
 XX
 DE Artificial sequence designed peptide #26.
 XX
 KM Peptide synthesis; chemical synthesis; solid phase synthesis.
 XX
 OS Synthetic.
 XX
 PN WO200055182-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 15-MAR-2000; 2000WO-0P01584.
 XX
 PR 15-MAR-1999; 99JP-0067917.
 XX
 PA (SANY) SANKYO CO LTD.
 XX
 PI Kawaguchi J, Serizawa N;
 XX
 DR WPI; 2000-602106/57.
 XX
 PT Chemically synthesizing a peptide by solid phase synthesis by adding
 PT protective group to inactivate part of alpha amino group of an amino
 PT acid group -
 XX
 PS Example 6; Page 23; 38pp; Japanese.
 XX
 CC The present invention describes a method for chemically synthesizing a
 CC peptide by solid phase synthesis. The method comprises processing a
 CC carrier resin to which the C-terminal amino acid derivative of the
 CC target peptide has been preliminarily bonded, is characterized by adding
 CC a protective group to inactivate a part of alpha-amino groups of an
 CC amino acid capable of undergoing a peptide extension reaction on the
 CC carrier resin. The method is useful for synthesizing a peptide by solid
 CC phase synthesis. Long chain peptides can be readily synthesised in with
 CC a high success ratio without using the segment condensation method or
 CC gene manipulations. The present sequence represents an amino acid
 CC peptide sequence given in an example from the present invention.

XX Sequence 74 AA;
SQ

Query Match 34.6%; Score 140.5; DB 21; Length 74;
Best Local Similarity 48.0%; Pred. No. 1.3e-11;
Matches 36; Conservative 2; Mismatches 8; Indels 29; Gaps 2;

OY 1 MKVTVAFNQGFENRRVFIKRVSNVTHGRRIDIPASKNPHLOKNTIGTRISLKTSGK 60
DB 6 MKVTVAFNQGF-----DIPASKNPHLOK-----KLTSGK 36
OY 61 IASRRVDGIIAAYON 75
DB 37 IASCLNMGVHVANN 51

RESULT 9

AAB23906
ID AAB23906 standard; peptide; 94 AA.

XX AAB23906;
AC

DT 17-JAN-2001 (first entry)

DE Artificial sequence designed peptide #27.

KW Peptide synthesis; chemical synthesis; solid phase synthesis.

OS Synthetic.

PN W0200055182-A1.

PD 21-SEP-2000.

PF 15-MAR-2000; 2000MO-JP01584.

PR 15-MAR-1999; 99JP-0067917.

PS (SANY) SANKYO CO LTD.

PI Kawaguchi J, Serizawa N;

DR WPI; 2000-602106/57.

PT Chemically synthesizing a peptide by solid phase synthesis by adding
PT protective group to inactivate part of alpha amino group of an amino
PT acid group

PS Example 6; Page 24; 38pp; Japanese.

CC The present invention describes a method for chemically synthesizing a
CC peptide by solid phase synthesis. The method comprises processing a
CC carrier resin to which the C-terminal amino acid derivative of the
CC target peptide has been preliminarily bonded is characterized by adding
CC a protective group to inactivate a part of alpha-amino groups of an
CC amino acid capable of undergoing a peptide extension reaction on the
CC carrier resin. The method is useful for synthesizing a peptide by solid
CC phase synthesis. Long chain peptides can be readily synthesized in with
CC a high success ratio without using the segment condensation method or
CC gene manipulations. The present sequence represents an amino acid
CC peptide sequence given in an example from the present invention.

XX Sequence 94 AA;

Query Match 34.6%; Score 140.5; DB 21; Length 94;
Best Local Similarity 48.0%; Pred. No. 1.7e-11;
Matches 36; Conservative 2; Mismatches 8; Indels 29; Gaps 2;

OY 1 MKVTVAFNQGFENRRVFIKRVSNVTHGRRIDIPASKNPHLOKNTIGTRISLKTSGK 60
DB 26 MKVTVAFNQGF-----DIPASKNPHLOK-----KLTSGK 56
OY 61 IASRRVDGIIAAYON 75

DB 57 IASCLNMGVHVANN 71

RESULT 10

AAB23897
ID AAB23897 standard; peptide; 95 AA.

XX AAB23897;
AC

DT 17-JAN-2001 (first entry)

DE Artificial sequence designed peptide SEQ ID NO:5.

KW Peptide synthesis; chemical synthesis; solid phase synthesis.

OS Synthetic.

PN W0200055182-A1.

PD 21-SEP-2000.

PF 15-MAR-2000; 2000MO-JP01584.

PR 15-MAR-1999; 99JP-0067917.

PS (SANY) SANKYO CO LTD.

PI Kawaguchi J, Serizawa N;

DR WPI; 2000-602106/57.

PT Chemically synthesizing a peptide by solid phase synthesis by adding
PT protective group to inactivate part of alpha amino group of an amino
PT acid group

PS Example 5; Page 34-35; 38pp; Japanese.

CC The present invention describes a method for chemically synthesizing a
CC peptide by solid phase synthesis. The method comprises processing a
CC carrier resin to which the C-terminal amino acid derivative of the
CC target peptide has been preliminarily bonded is characterized by adding
CC a protective group to inactivate a part of alpha-amino groups of an
CC amino acid capable of undergoing a peptide extension reaction on the
CC carrier resin. The method is useful for synthesizing a peptide by solid
CC phase synthesis. Long chain peptides can be readily synthesized in with
CC a high success ratio without using the segment condensation method or
CC gene manipulations. The present sequence represents an amino acid
CC peptide sequence given in an example from the present invention.

XX Sequence 95 AA;

Query Match 34.6%; Score 140.5; DB 21; Length 95;
Best Local Similarity 48.0%; Pred. No. 1.7e-11;
Matches 36; Conservative 2; Mismatches 8; Indels 29; Gaps 2;

OY 1 MKVTVAFNQGFENRRVFIKRVSNVTHGRRIDIPASKNPHLOKNTIGTRISLKTSGK 60
DB 27 MKVTVAFNQGF-----DIPASKNPHLOK-----KLTSGK 57
OY 61 IASRRVDGIIAAYON 75
DB 58 IASCLNMGVHVANN 72

RESULT 11

AAB23901
ID AAB23901 standard; peptide; 95 AA.

XX AAB23901;
AC

DT 17-JAN-2001 (first entry)

DE Artificial sequence designed peptide #23.
 XX Peptide synthesis; chemical synthesis; solid phase synthesis.
 KW Synthetic.
 OS
 XX WO200055182-A1.
 PN
 XX 21-SEP-2000.
 PD
 XX 15-MAR-2000; 2000WO-JP01584.
 PF
 XX 15-MAR-1999; 99JP-0067917.
 PR
 XX (SANY) SANKYO CO LTD.
 PA
 XX Kawaguchi J, Serizawa N;
 PI
 XX WPI; 2000-602106/57.
 DR
 XX Chemically synthesizing a peptide by solid phase synthesis by adding
 PT protective group to inactivate part of alpha amino group of an amino
 CC acid group
 XX
 XX Example 5; Page 21-22; 38pp; Japanese.
 PS
 XX The present invention describes a method for chemically synthesizing a
 CC peptide by solid phase synthesis. The method comprises processing a
 CC carrier resin to which the C-terminal amino acid derivative of the
 CC target peptide has been preliminarily bonded is characterised by adding
 CC a protective group to inactivate a part of alpha-amino groups of an
 CC amino acid capable of undergoing a peptide extension reaction on the
 CC carrier resin. The method is useful for synthesizing a peptide by solid
 CC phase synthesis. Long chain peptides can be readily synthesised in with
 CC a high success ratio without using the segment condensation method or
 CC gene manipulations. The present sequence represents an amino acid
 CC peptide sequence given in an example from the present invention.
 CC
 SQ Sequence 95 AA;
 Query Match 34.6%; Score 140.5; DB 21; Length 95;
 Best Local Similarity 48.0%; Pred. No. 1.7e-11;
 Matches 36; Conservative 2; Mismatches 8; Indels 29; Gaps 2;
 QY 1 MKYTVAFNPGPNRRVFIKRVSNVITHGRRIDIFASKNFHLQKNTIGTRISLKTSGK 60
 DB 27 MKYTVAFNPGP-----DIFASKNFHLQKN-----KLTSGK 57
 QY 61 IASRRVDGITAAAYON 75
 DB 58 IASCLNMGVLHVANN 72
 RESULT 12
 AAB69098
 ID AAB69098 standard; Peptide; 95 AA.
 XX
 AC AAB69098;
 XX
 DT 23-APR-2001 (first entry)
 XX
 DE Cedar pollen allergen T cell epitope derived peptide SEQ ID NO:5.
 XX
 KW Japanese cedar; Cryptomeria japonica; cedar pollen allergen;
 KM T cell epitope; antiIgipollinosis.
 XX
 OS Cryptomeria japonica.
 OS Synthetic.
 XX
 PN JP2000327699-A.
 XX
 PD 28-NOV-2000.

PF 15-MAR-2000; 2000JP-0071710.
 XX
 PR 15-MAR-1999; 99JP-0068316.
 XX
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA (SANY) SANKYO CO LTD.
 XX
 DR WPI; 2001-185061/19.
 XX
 PT Novel peptide and its use -
 XX
 PS Claim 5; Page 33-34; 75pp; Japanese.
 CC
 CC The present invention describes a peptide, its complex, derivative or
 CC its polymerizate, where the peptide (I) has a formula of:
 CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7
 CC where alpha1 to alpha7 = amino acid sequence selected from amino acid
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes
 CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen
 CC allergens. The peptide can be used in antiIgipollinosis agent.
 CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used
 CC in the exemplification of the present invention.
 CC
 SQ Sequence 95 AA;
 Query Match 34.6%; Score 140.5; DB 22; Length 95;
 Best Local Similarity 48.0%; Pred. No. 1.7e-11;
 Matches 36; Conservative 2; Mismatches 8; Indels 29; Gaps 2;
 QY 1 MKYTVAFNPGPNRRVFIKRVSNVITHGRRIDIFASKNFHLQKNTIGTRISLKTSGK 60
 DB 27 MKYTVAFNPGP-----DIFASKNFHLQKN-----KLTSGK 57
 QY 61 IASRRVDGITAAAYON 75
 DB 58 IASCLNMGVLHVANN 72
 RESULT 13
 AAB69104
 ID AAB69104 standard; Protein; 96 AA.
 XX
 AC AAB69104;
 XX
 DT 23-APR-2001 (first entry)
 XX
 DE Cedar pollen allergen T cell epitope derived protein SEQ ID NO:13.
 XX
 KW Japanese cedar; Cryptomeria japonica; cedar pollen allergen;
 KM T cell epitope; antiIgipollinosis.
 XX
 OS Cryptomeria japonica.
 OS Synthetic.
 XX
 PN JP2000327699-A.
 XX
 PD 28-NOV-2000.
 XX
 PF 15-MAR-2000; 2000JP-0071710.
 XX
 PR 15-MAR-1999; 99JP-0068316.
 XX
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA (SANY) SANKYO CO LTD.
 XX
 DR WPI; 2001-185061/19.
 XX
 PT Novel peptide and its use -
 XX
 PS Claim 6; Page 40; 75pp; Japanese.
 CC
 CC The present invention describes a peptide, its complex, derivative or
 CC its polymerizate, where the peptide (I) has a formula of:

CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7
 CC where alpha1 to alpha7 = amino acid sequence selected from amino acid
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes
 CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen
 CC allergens. The peptide can be used in an anti-sugipollinosis agent.
 CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used
 CC in the exemplification of the present invention.
 XX

Sequence 96 AA;

Query Match 34.6%; Score 140.5; DB 22; Length 96;
 Best Local Similarity 48.0%; Pred. No. 1.8e-11;
 Matches 36; Conservative 2; Mismatches 8; Indels 29; Gaps 2;

QY 1 MKVTVAFNQGFENRRVFIKRVSNVHIGRRIDIFASKNPHLOKNTIGTGRRISLKTSGK 60
 DB 15 MKVTVAFNQGF-----DIFASKNPHLOK-----KLTSGK 45
 QY 61 IASRRVDGIIAAYON 75
 DB 46 IASCLNTGLVHVANN 60

RESULT 14
 AAB69119
 ID AAB69119 standard; Protein; 96 AA.
 XX
 AC AAB69119;
 XX
 DT 23-APR-2001 (first entry)
 XX

DE Cedar pollen allergen T cell epitope derived protein SEQ ID NO:53.

KW Japanese cedar; Cryptomeria japonica; cedar pollen allergen;
 KM T cell epitope; anti-sugipollinosis.

XX Cryptomeria japonica.
 OS Synthetic.
 OS
 XX JP2000327699-A.
 PN
 XX
 XX 28-NOV-2000.
 PD
 PF 15-MAR-2000; 2000JP-0071710.
 XX
 XX 15-MAR-1999; 99JP-0068316.
 PR
 XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA (SANY) SANKYO CO LTD.
 XX
 XX WPI; 2001-185061/19.
 DR N-PSDB; AAF59037.
 XX
 PT Novel peptide and its use -
 XX
 XX Example 11; Page 54-55; 75pp; Japanese.
 PS
 XX

CC The present invention describes a peptide, its complex, derivative or
 CC its polymerizate, where the peptide (1) has a formula of:
 CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7
 CC where alpha1 to alpha7 = amino acid sequence selected from amino acid
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes
 CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen
 CC allergens. The peptide can be used in an anti-sugipollinosis agent.
 CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used
 CC in the exemplification of the present invention.
 XX

Sequence 96 AA;

Query Match 34.6%; Score 140.5; DB 22; Length 96;
 Best Local Similarity 48.0%; Pred. No. 1.8e-11;
 Matches 36; Conservative 2; Mismatches 8; Indels 29; Gaps 2;

QY 1 MKVTVAFNQGFENRRVFIKRVSNVHIGRRIDIFASKNPHLOKNTIGTGRRISLKTSGK 60
 DB 15 MKVTVAFNQGF-----DIFASKNPHLOK-----KLTSGK 45
 QY 61 IASRRVDGIIAAYON 75
 DB 46 IASCLNTGLVHVANN 60

RESULT 15
 AAB69105
 ID AAB69105 standard; Protein; 97 AA.
 XX
 AC AAB69105;
 XX
 DT 23-APR-2001 (first entry)
 XX

DE Cedar pollen allergen T cell epitope derived protein SEQ ID NO:15.

KW Japanese cedar; Cryptomeria japonica; cedar pollen allergen;
 KM T cell epitope; anti-sugipollinosis.

XX Cryptomeria japonica.
 OS Synthetic.
 OS
 XX JP2000327699-A.
 PN
 XX
 XX 28-NOV-2000.
 PD
 PF 15-MAR-2000; 2000JP-0071710.
 XX
 XX 15-MAR-1999; 99JP-0068316.
 PR
 XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA (SANY) SANKYO CO LTD.
 XX
 XX WPI; 2001-185061/19.
 DR N-PSDB; AAF59013.
 XX
 PT Novel peptide and its use -
 XX
 XX Claim 9; Page 41-42; 75pp; Japanese.
 PS
 XX

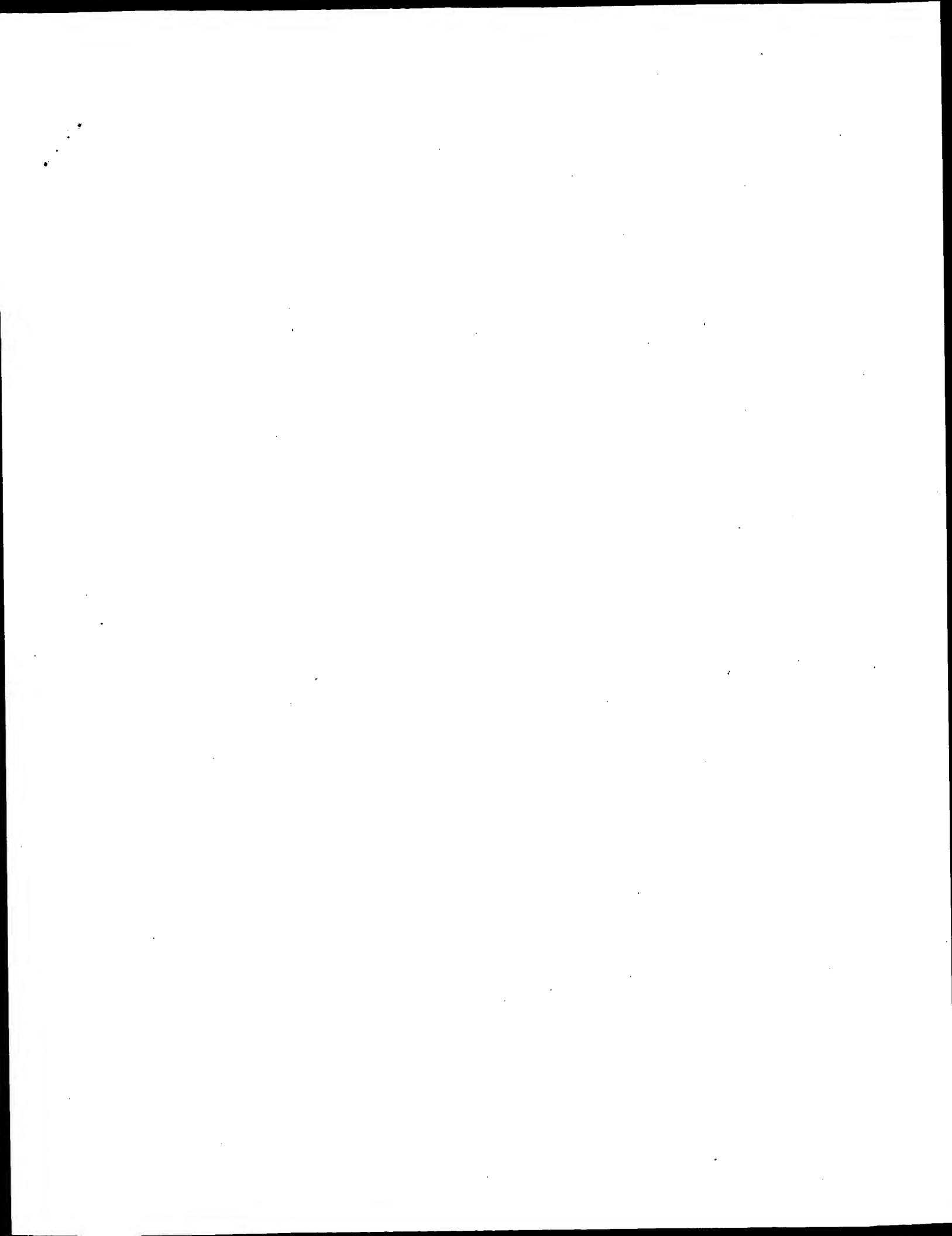
CC The present invention describes a peptide, its complex, derivative or
 CC its polymerizate, where the peptide (1) has a formula of:
 CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7
 CC where alpha1 to alpha7 = amino acid sequence selected from amino acid
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes
 CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen
 CC allergens. The peptide can be used in an anti-sugipollinosis agent.
 CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used
 CC in the exemplification of the present invention.
 XX

Sequence 97 AA;

Query Match 34.6%; Score 140.5; DB 22; Length 97;
 Best Local Similarity 48.0%; Pred. No. 1.8e-11;
 Matches 36; Conservative 2; Mismatches 8; Indels 29; Gaps 2;

QY 1 MKVTVAFNQGFENRRVFIKRVSNVHIGRRIDIFASKNPHLOKNTIGTGRRISLKTSGK 60
 DB 16 MKVTVAFNQGF-----DIFASKNPHLOK-----KLTSGK 46
 QY 61 IASRRVDGIIAAYON 75
 DB 47 IASCLNTGLVHVANN 61

Search completed: April 20, 2003, 13:06:02
 Job time : 98.0526 secs



GenCore version 5.1.4 ps 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:57:30 / Search time 42.5526 Seconds
(without alignments)
237.215 Million cell updates/sec

Title: US-09-142-524D-2

Perfect score: 558
Sequence: 1 MKVTVAENQFGPNRRVFIKR.....FDGRVDGIIAAYQNPASWK 105

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR73:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115.5	20.7	514	2	S48730 Cry j II protein -
2	115	20.6	514	2	second major aller
3	98	17.6	514	2	polygalacturonase
4	93.5	16.8	374	2	JC2124 major allergen Cry
5	93.5	16.8	374	2	JC2123 major allergen Cry
6	91.5	16.4	507	2	JC7366 Jun a 2 protein -
7	84.5	15.1	398	2	F72335 hypothetical prote
8	81	14.5	219	2	C84647 probable synaptoch
9	77	13.8	219	2	T04630 synaptoch
10	75.5	13.5	477	2	S76496 hypothetical prote
11	73.5	13.2	644	2	AD3471 adenyl-1-sulfate k
12	71	12.7	576	1	AA5049 4-alpha-glucanotra
13	68.5	12.3	390	2	B57154 phosphopentomutase
14	67.5	12.1	1070	2	S46755 hypothetical prote
15	67	12.0	689	1	JW0107 very-long-chain ac
16	66.5	11.9	110	2	D48562 NADH2 dehydrogenas
17	66.5	11.9	777	2	T13561 hypothetical prote
18	66.5	11.8	777	2	G69310 penicillin G acyla
19	66	11.8	188	2	B72450 probable membrane
20	66	11.8	196	2	S54580 hypothetical prote
21	65.5	11.7	493	2	AB3032 exon protein (A722
22	65.5	11.7	602	2	A96254 oxidoreductases ho
23	64.5	11.6	317	2	AD1542 hypothetical 22k p
24	64	11.5	284	2	JC2330 major antigenic pr
25	64	11.5	350	2	S75065 sensory transducti
26	64	11.5	612	2	H81043 hypothetical prote
27	64	11.5	612	2	T13754 NADH2 dehydrogenas
28	64	11.5	612	2	T13754 NADH2 dehydrogenas
29	64	11.5	3194	2	D71917 toxin-like outer m

ALIGNMENTS

RESULT 1

S48730

Cry j II protein - Japanese cedar

C/Species: Cryptomeria japonica (Japanese cedar)

C/Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 28-May-1999

C/Accession: S48730

R/Namba, M.; Kurose, M.; Toriigo, K.; Hino, K.; Taniguchi, Y.; Fukuda, S.; Usui, M.; Kuri

FEBS Lett. 353, 124-128, 1994

A/Title: Molecular cloning of the second major allergen, Cry j II, from Japanese cedar p

A/Reference number: S48730, MUID:95010777, PMID:7926035

A/Accession: S48730

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-514 <N>M>

A/Cross-references: GB:D37765; NID:G577695; PIDN:BA07021.1; PID:0007598; PID:G577696

Query Match

Best Local Similarity 20.7%; Score 115.5; DB 2; Length 514;

Matches 29; Conservativity 9; Mismatches 11; Indels 21; Gaps 4;

QY

12 PNRVFIKRVSNVLIHGR-----RID--IFAKSKFHIQKNTIGGRKNNRIWLOQA 62

Db 96 PNMKKFV--VNNLFNFGPOCPHFTKVDGIIAAYQN-----PASKNRRITWLOPA 143

QY 63 KLTGFTLMGR 72

Db 144 KLTGFTLMGR 153

RESULT 2

JC2498

second major allergen Cry j II precursor - Japanese cedar

C/Species: Cryptomeria japonica (Japanese cedar)

C/Date: 16-Mar-1995 #sequence_revision 26-May-1995 #text_change 21-Jul-2000

C/Accession: JC2498; PC2346; A60147

R/Komiyama, N.; Some, T.; Shimizu, K.; Morikubo, K.; Kino, K.

Biochem. Biophys. Res. Commun. 201, 1021-1028, 1994

A/Title: cDNA cloning and expression of Cry j II, the second major allergen of Japanese c

A/Reference number: JC2498; MUID:94271186; PMID:8002972

A/Accession: JC2498

A/Molecule type: mRNA

A/Residues: 1-514 <K>M>

A/Cross-references: DDBJ:D29772; NID:G506857; PIDN:BA06172.1; PID:G506858

A/Accession: PC2346

A/Molecule type: protein

A/Residues: 52-61 <K>O2>

R/Sakaguchi, M.; Inoue, S.; Tanai, M.; Ando, S.; Usui, M.; Matuhasi, T.

Allergy 45, 309-312, 1990

A/Title: Identification of the second major allergen of Japanese cedar pollen.

A/Reference number: A60147; MUID:90342988; PMID:2382797

A/Accession: A60147

A/Molecule type: protein

A:Residues: 55-64 <SAK>

C:Keywords: glycoprotein; pollen

F:1-54/Domain: signal sequence #status predicted <SIG>

F:55-460/Product: second major allergen Cry j #status predicted <MAT>

F:429,460/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

20.6%; Score 115; DB 2; Length 514;

Best Local Similarity 95.2%; Pred. No. 2,5e-05;

Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 52 WKNNRIWLOFAKLTGFTLMGR 72

DB 133 WKNNRIWLOFAKLTGFTLMGR 153

RESULT 3

polygalacturonase Cha o 2 - Japanese cypress

C:Species: Chamaecyparis obtusa (Japanese cypress)

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000

C:Accession: J07100; PC7026

R:Morii, T.; Yokoyama, M.; Komiyama, N.; Okano, M.; Kino, K.

Biochem. Biophys. Res. Commun. 263: 166-171, 1999

A:Title: Purification, identification, and cDNA cloning of Cha o 2, the second major all

A:Reference number: J07100; MUID:99417540; PMID:10486272

A:Accession: J07100

A:Molecule type: mRNA

A:Residues: 1-514 <MOR>

A:Accession: PC7026

A:Molecule type: protein

A:Residues: 51-62 <MO2>

Query Match

17.6%; Score 98; DB 2; Length 514;

Best Local Similarity 38.5%; Pred. No. 0.0023;

Matches 25; Conservative 8; Mismatches 16; Indels 16; Gaps 3;

QY 13 NRRVPIKRVSNVTHGRIDIFASKNFHLQKNTGT-----GRRWKNNRIWLOFAKLTG 66

DB 98 NKKFF---VNNLVFRG-----PCQPHLSFKVDGTIVADPAPAKSKKIWLOFAQLTD 147

QY 67 FTLMG 71

DB 148 FTLMG 152

RESULT 4

major allergen Cry j I precursor (clone pCCI-15) - Japanese cedar

C:Species: Cryptomeria japonica (Japanese cedar)

C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 21-Jul-2000

C:Accession: J02124

R:Some, T.; Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.

Biochem. Biophys. Res. Commun. 199, 619-625, 1994

A:Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japanese

A:Reference number: J02123; MUID:94183234; PMID:8135802

A:Accession: J02124

A:Molecule type: mRNA

A:Residues: 1-374 <SON>

A:Cross-references: GB:D26545; NID:g493631; PIDN:BA05543.1; PID:g493634

A:Experimental source: pollen

A:Note: the authors described carbohydrate binding site for residue 279

C:Superfamily: pectate lyase LAY59

C:Keywords: glycoprotein; pollen

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-374/Product: major allergen Cry j I (clone pCCI-15) #status predicted <MAT>

F:158,191,293,354/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

16.8%; Score 93.5; DB 2; Length 374;

Best Local Similarity 51.4%; Pred. No. 0.0054;

Matches 19; Conservative 6; Mismatches 7; Indels 5; Gaps 1;

QY 53 KNNRIWLOFAKLTGFTLMGRRLKMPYIAGYKTFDGR 89

DB 113 KNNRIWLOFAKLTGFTLMGRRLKMPYIAGYKTFDGR 89

DB 85 RDRPLMIIFS-----GNMNIKLKMPYIAGYKTFDGR 116

RESULT 5

major allergen Cry j I precursor (clone pCCI-2-2) - Japanese cedar

C:Species: Cryptomeria japonica (Japanese cedar)

C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 21-Jul-2000

C:Accession: J02123; PC2065

R:Some, T.; Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.

Biochem. Biophys. Res. Commun. 199, 619-625, 1994

A:Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japanese

A:Reference number: J02123; MUID:94183234; PMID:8135802

A:Accession: J02123

A:Molecule type: mRNA

A:Residues: 1-374 <SON>

A:Cross-references: GB:D26544; NID:g493631; PIDN:BA05542.1; PID:g493632

A:Experimental source: pollen

A:Accession: PC2065

A:Molecule type: protein

A:Residues: 22-53;58-81;219-232;236-258;299-307;346-372 <SO2>

A:Note: the authors described carbohydrate binding site for residue 279

C:Superfamily: pectate lyase LAY59

C:Keywords: glycoprotein; pollen

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-374/Product: major allergen Cry j I (clone pCCI-2-2) #status predicted <MAT>

F:158,191,293,354/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

16.8%; Score 93.5; DB 2; Length 374;

Best Local Similarity 51.4%; Pred. No. 0.0054;

Matches 19; Conservative 6; Mismatches 7; Indels 5; Gaps 1;

QY 53 KNNRIWLOFAKLTGFTLMGRRLKMPYIAGYKTFDGR 89

DB 85 RDRPLMIIFS-----GNMNIKLKMPYIAGYKTFDGR 116

RESULT 6

Jun a 2 protein - mountain cedar

C:Species: Juniperus ashei (mountain cedar)

C:Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000

C:Accession: J07366; PC7093

R:Yokoyama, M.; Miyahara, M.; Shimizu, K.; Kino, K.; Tsunoo, H.

Biochem. Biophys. Res. Commun. 275, 195-202, 2000

A:Title: Purification, identification, and cDNA cloning of Jun a 2, the second major alle

A:Reference number: J07366

A:Accession: J07366

A:Molecule type: mRNA

A:Residues: 1-507 <YOK>

A:Cross-references: GB:AJ404653

A:Accession: PC7093

A:Molecule type: protein

A:Residues: 55-63 <YO2>

C:Comment: This protein, a second major allergen of mountain cedar pollen, which is invo

C:Keywords: glycoprotein; pollen

Query Match

16.4%; Score 91.5; DB 2; Length 507;

Best Local Similarity 31.5%; Pred. No. 0.013; 26; Indels 23; Gaps 5;

Matches 29; Conservative 14; Mismatches 26; Indels 23; Gaps 5;

QY 12 PNRVPIKRVSNVTHGRIDIFASKNFHLQKNTGT---GRRWKNNRIWLOFAKLTGFT 68

DB 97 PANKKFV---VNNLVFRG---PCQPHLSFKVDGTIVADPAPAKSKKIWLOFAQLTD 150

QY 69 LMGRLKMPYIAGY-----KTFDGRRV 91

DB 151 LMGTV-----IDGGNRWMSDQCKTINGRIV 177

RESULT 7

F72335

hypothetical protein TM0760 - Thermotoga maritima (strain MSB8)

C:Species: Thermotoga maritima

C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000

C/Accession: F72335

R: Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq

A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Accession: F72335

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-398 <ARN>

A:Cross-references: GB:AE001746; GB:AE000512; NID:94981285; PIDN:AA035842.1; PID:9498128

C:Genetics:

A:Gene: TM0760

Query Match

Best Local Similarity 15.1%; Score 8445; DB 2; Length 398;

Matches 25; Conservative 11; Mismatches 35; Indels 17; Gaps 2;

28 GRRIDFASKNFHLOKNTTGTGRKNNRIW-----LQPAKLTGFTLMGRRLKM 76

33 GHRVDIYIDFSL-----TGRWSETFQWSPSKDGVPIVEREYTGNSLSLSI 86

77 PMVIAGYTFDGRVDGIIAAYONPASM 104

87 DYRNCRKTIQKRVVITASSHPSP 114

RESULT 8

C84647

probable synaprobrevin (imported) - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C/Accession: G64647

R: Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon, L.; Neus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A64420; MUID:20083487; PMID:10617197

A:Accession: G64647

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-219 <STO>

A:Cross-references: GB:AE002093; NID:94567243; PIDN:AMD23657.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g25340

A:Map position: 2

Query Match

Best Local Similarity 14.5%; Score 81; DB 2; Length 219;

Matches 16; Conservative 14; Mismatches 22; Indels 2; Gaps 1;

14 RRVFKRVSVNIIHGRRIDFASKNFHLOKNTTGTGRKNNRIW-----LQPAKLTGFTLMGRRLKM 65

138 RDVMIENIDNLDGRERELVDKTNAMQNTFRKQTRFPNNTVMWKNCKLT 191

RESULT 9

T04630

synaprobrevin homolog F10N7_40 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-May-1999

C/Accession: T04630

R: Beyan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Hohnsbeil, J.; Mewes, H.W.; Mayer, K.F.X

submitted to the Protein Sequence Database, March 1999

A:Reference number: Z15263

A:Accession: T04630

A:Molecule type: DNA

A:Residues: 1-219 <BRV>

A:Cross-references: EMBL:AL021636

A:Experimental source: cultivar Columbia; BAC clone F10N7

C:Genetics:

A:Map position: 4

A:Introns: 73/1; 136/3; 191/2

A>Note: F10N7_40

Query Match

Best Local Similarity 13.8%; Score 77; DB 2; Length 219;

Matches 22; Conservative 16; Mismatches 27; Indels 6; Gaps 3;

14 RRVFKRVSVNIIHGRRIDFASKNFHLOKNTTGTGRKNNRIW-----LQPAKLTGFTLMGRRLKM 70

138 RGVMIENIDNLDGRERELVDKTNAMQNTFRKQTRFPNNTVMWKNCKLT--VLL 194

71 GRRLKMPVIA 81

195 ILLLVITIA 205

RESULT 10

S76496

hypothetical protein - Synechocystis sp. (strain PCC 6803)

C:Species: Synechocystis sp.

A:Variety: PCC 6803

C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000

C/Accession: S76496

R: Kaneke, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S76496

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-477 <KAN>

A:Cross-references: EMBL:D90915; GB:AB001339; NID:91653604; PIDN:BA18625.1; PID:9165371

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Superfamily: starch synthase

Query Match

Best Local Similarity 13.5%; Score 75.5; DB 2; Length 477;

Matches 23; Conservative 15; Mismatches 29; Indels 19; Gaps 4;

8 NQFQNRVYFKRVSVNIIHGRRIDFASKNFHLOKNTTGTGRKNNRIW-----LQPAKLTGFTLMGRRLKM 67

31 HQGHVDFVRFMPYGPFI---GDKIDV-----EKRPV---WKGRAMFOQFAVQSY 74

68 TLMGRRLKMPVITAGYKTFDGRVDG 93

75 L--PDTKIPLYLFGHPADSRRIYG 97

RESULT 11

AD3471

adenylate sulfatase (EC 2.7.1.25) (imported) - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis

C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 19-Apr-2002

C/Accession: AD3471

R: DelVecchio, V.G.; Kaparat, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, I.; Mazur, M.; Goldstein, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Latesac, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A:Reference number: AD3252; PMID:11756688

A:Accession: AD3471

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-644 <KUD>

A:Cross-references: GB:AE009917; PIDN:AAU52935.1; PID:917983784; GSPDB:GN00190

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BME11754
A:Map position: 1
C:Superfamily: nucleation protein nod; adenyllyl sulfate kinase homology; translation etc
C:Keywords: phosphotransferase

Query Match 13.2%; Score 73.5; DB 2; Length 644;
Best Local Similarity 26.0%; Pred. No. 2.1;
Matches 26; Conservative 17; Mismatches 44; Indels 13; Gaps 4;

5 VAFNPGFNR---VFIRVSVNIIHGRID--IFASKNFHLO-----KNTIGTGRMK 53
Db 402 IAFDYGERRATGAFVLDRLNATVGMGMDALRQATVNHLOAFDLNQARAAGKFGK 461
Qy 54 NNRWLOPAKLTGFTLMGRRLKMPYIAGYKTF--DGRV 91
Db 462 PAVLWFTGASGKSTIANRLRQALHAGKHTYLLDGDNV 501

RESULT 12

A45049
A:alpha-glucanotransferase (EC 2.4.1.25) - potato
N:Alternate names: amylomalase; D-enzyme; disproportionating enzyme
C:Species: Solanum tuberosum (potato)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A45049
R:Takaha, T.; Yanase, M.; Okada, S.; Smith, S.M.
J. Biol. Chem. 268, 1391-1396, 1993
A>Title: Disproportionating enzyme (4-alpha-glucanotransferase; EC 2.4.1.25) of potato.
A:Reference number: A45049; MUID:93123262; PMID:7678257
A:Accession: A45049
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-576 <TAK>
A:Cross-references: GB:X68664; GB:S52648; NID:9296691; PIDN:CAA48630.1; PID:9296692
A:Experimental source: cv. May Queen, tubers
A>Note: sequence extracted from NCBI backbone (NCBI:122077, NCBI:P.122078)
C:Superfamily: 4-alpha-glucanotransferase
C:Keywords: chloroplast; glycosyltransferase; hexosyltransferase

Query Match 12.7%; Score 71; DB 1; Length 576;
Best Local Similarity 35.1%; Pred. No. 3.6;
Matches 20; Conservative 10; Mismatches 13; Indels 14; Gaps 4;

Qy 31 IIFPSKYNHLQKNTIGTGRMKNNRIMLOPAKLTGFTLMGRRLKMPYIAGYKTFD 87
Db 262 IDIFIAQQLFQ-----RQMKVYR---DYARSKGISIMG---DMPYV-GYHSAD 304

RESULT 13

E97154
A:phosphotransferase [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C:Accession: E97154
R:Noiling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A>Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: E97154
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-350 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK80024.1; PID:g15025052; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC2065
C:Superfamily: phosphotransferase

Query Match 12.3%; Score 68.5; DB 2; Length 390;
Best Local Similarity 29.0%; Pred. No. 4.5;
Matches 27; Conservative 12; Mismatches 45; Indels 9; Gaps 4;

Qy 6 AFNPGFNRVFIK-RVSVNIIHGRIDIPASKN----FHLQKNTIGTGR--WKNNRI 57
Db 213 AADPPNKTMLLEYIKNNKSNVNCVGGIEDIFNKKVTEAVHILKNNMDGVDKLEYKTDKK 272
Qy 58 WLOPAKLTGFTLM--GRRLKMPYIAGYKTFDGR 89
Db 273 GLITNLVDFTMLYGRNDPKGYAKALEFDFGR 305

RESULT 14

S46755
A:hypothetical protein YHR154w - Yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C>Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 19-Apr-2002
C:Accession: S46755
R:Macri, C.
Submitted to the EMBL Data Library, June 1994
A:Description: The sequence of S. cerevisiae cosmid 9666.
A:Reference number: S46752
A:Accession: S46755
A:Molecule type: DNA
A:Residues: 1-1070 <MAC>
A:Cross-references: EMBL:U10397; NID:9500647; PID:9500651; GSPDB:GN00008; MIPS:YHR154w
A:Gene: ESC4; MIPS:YHR154w
A:Cross-references: SCD:S0001197
A:Map position: 8R

Query Match 12.1%; Score 67.5; DB 2; Length 1070;
Best Local Similarity 27.4%; Pred. No. 18;
Matches 20; Conservative 15; Mismatches 23; Indels 15; Gaps 4;

Qy 2 KTVAFNPGFNRVFIKRVSVNIIHGRIDIPASKNHL-QKNTIG-----TGRWK--- 53
Db 379 ELTVATVNFVSGRFYQRLV-ILGSLTPELTRKNTLITKSTIGKKFVAKKWSLDP 437
Qy 54 -----NNRIMLO 60
Db 438 QNAILVTNHWML 450

RESULT 15

JM0107
A:very-long-chain acyl-CoA synthetase related protein - mouse
N:Alternate names: VLACSR
C:Species: Mus musculus (house mouse)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 08-Sep-2000
C:Accession: JM0107
R:Berger, J.; Truppe, C.; Neumann, H.; Forss-Peterson, S.
Biochem. Biophys. Res. Commun. 247, 255-260, 1998
A>Title: A novel relative of the very-long-chain acyl-CoA synthetase and fatty acid trans-
A:Reference number: JM0107; MUID:98308102; PMID:9642112
A:Accession: JM0107
A:Molecule type: mRNA
A:Residues: 1-689 <BER>
A:Cross-references: GB:A1223959
A:Experimental source: liver
C:Comment: this protein likely functions as a plasma membrane transporter of long chain
cids.
F:169-647/Domain: acetate-CoA ligase homology <ACL>

Query Match 12.0%; Score 67; DB 1; Length 689;
Best Local Similarity 21.7%; Pred. No. 13;
Matches 25; Conservative 24; Mismatches 42; Indels 24; Gaps 4;

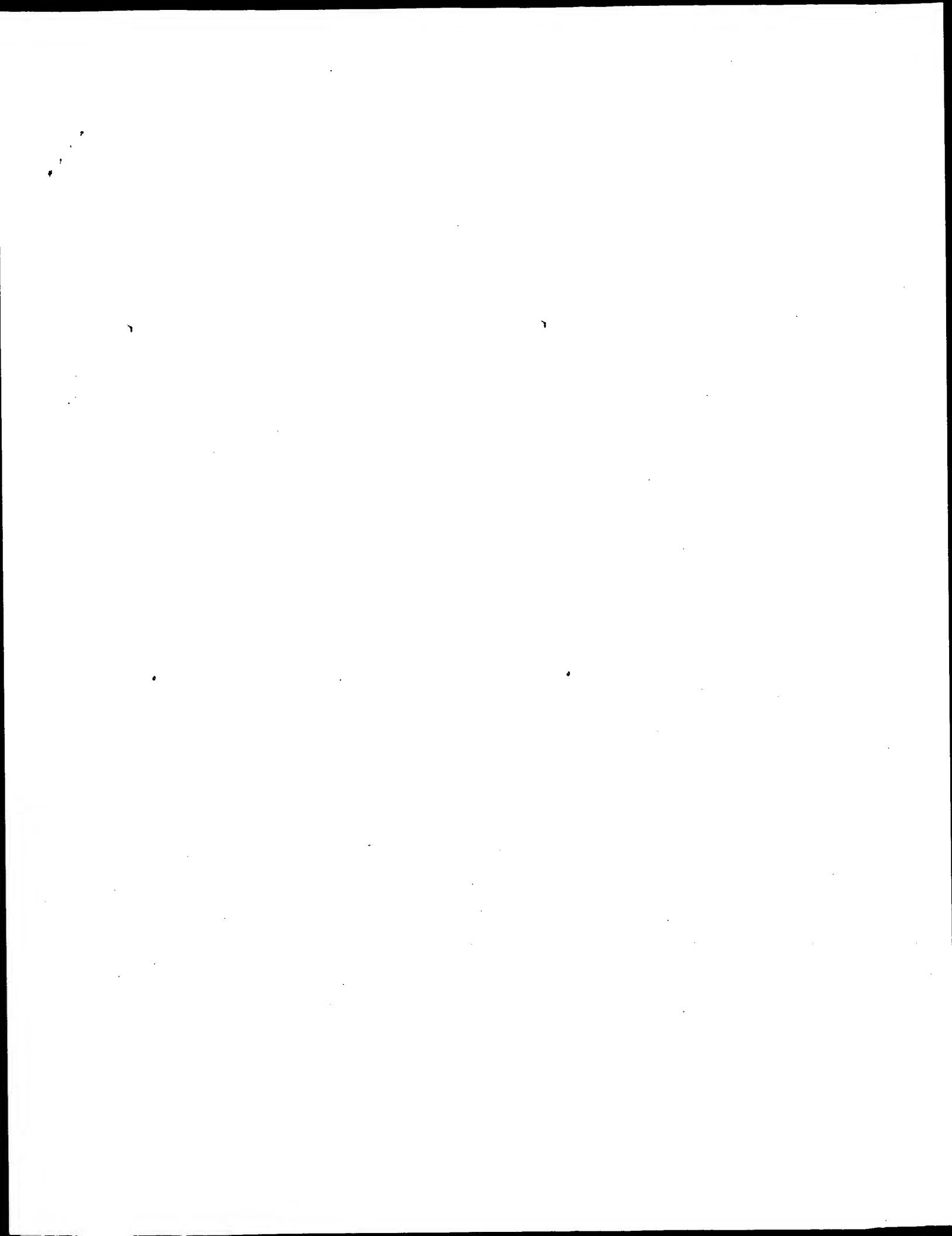
Qy 8 NQFPGNRVR--IKRVSVNIIHGRIDIPASKNFHLQKNTIGTGRKNNRI----- 57
Db 510 SQAESNRKLVANRRVGDLYFNNGDVLTDQSGFFYFQDRLDTPFRKGENVSTGEVSU 569
Qy 58 -----WLOPAKLTGFTLMGRRLKMPYIAGYKTFDGRVDDGTAAYQNPASW 104
Db 570 LSSIDFLFENVNVGVFVPGCGEGKVMGAVALAPGKTFDQKL-----YQHVRSW 618

Wed Apr 23 18:08:51 2003

us-09-142-524d-2.rpr

Page 5

Search completed: April 20, 2003, 13:15:37
Job time : 45.5526 secs



GenCore version 5.1.4 ps 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:25 ; Search time 20.4474 Seconds

(without alignments)
212.987 Million cell updates/sec

Title: US-09-142-524D-2

Perfect score: 558

Sequence: 1 MKVTVAENQFGPNRRVFIKR.....FDGRRVDGIIAAYQNPASWK 105

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	115.5	20.7	514	MPA2_CRYVA	P43212 cryptomeria
2	93.5	16.8	374	SBP_CRYVA	P18632 cryptomeria
3	85.5	15.3	367	MPA1_TJNAS	P81294 juniperus a
4	82.5	14.8	346	MPA1_CUPAK	Q98C99 cupressus a
5	75.5	13.5	477	GLGA_SYNY3	P74521 synechocyst
6	74.5	13.4	375	MPA1_CHAOB	Q96385 chamaecypar
7	71	12.7	576	DEEP_SOLTU	Q06801 solanum tub
8	68.5	12.3	234	A29B_DROME	Q46197 drosophila
9	67.5	12.1	1070	YHVA_YEAST	P38850 saccharomyc
10	66.5	11.9	110	Y12K_YEAST	P36289 san miguel
11	66	11.8	196	YMO7_YEAST	Q04487 saccharomyc
12	65.5	11.7	493	AMVR_DROSE	Q76261 drosophila
13	65.5	11.7	699	NUSC_DIGOR	Q32131 digitalis g
14	64.5	11.6	493	AMVR_DROOR	Q77015 drosophila
15	63.5	11.4	130	AMVR_DROOR	Q93459 chaiseira m
16	63.5	11.4	540	LAC3_THACU	Q02079 charatephor
17	63	11.3	572	Z136_HUMAN	P54737 homo sapien
18	63	11.3	880	Z136_HUMAN	P54737 homo sapien
19	62.5	11.2	449	YMD5_CABEL	P54737 homo sapien
20	62.5	11.2	493	PE59_LYCES	P15772 lycopersico
21	62.5	11.2	493	AMVR_DROBE	Q93459 chaiseira m
22	62.5	11.2	493	AMVR_DROME	Q77015 drosophila
23	62.5	11.2	493	AMVR_DROTE	Q77015 drosophila
24	62.5	11.2	493	AMVR_DROTE	Q77015 drosophila
25	62	11.1	493	AMVR_DROTE	Q77015 drosophila
26	61.5	11.0	493	AMVR_DROTE	Q77015 drosophila
27	61.5	11.0	493	AMVR_DROTE	Q77015 drosophila
28	61.5	11.0	493	AMVR_DROTE	Q77015 drosophila
29	61.5	11.0	493	AMVR_DROTE	Q77015 drosophila
30	61	10.9	275	PA5_FIG	Q94966 gallus gall
31	61	10.9	275	PA5_FIG	Q94966 gallus gall
32	60.5	10.8	505	VNS2_DSDNV	Q71154 diatraea sa
33	60.5	10.8	633	CSA5_DROME	Q07309 r nodg biflu
			706	KPCT_HUMAN	Q04759 homo sapien

34	60.5	10.8	1010	1	CLPP_CHLEU	P42379 chlamydomon
35	60.5	10.8	2133	1	PA8_FIG	P12263 sus scrofa
36	60	10.8	130	1	RS9_XYLEFA	Q99443 xylella fas
37	60	10.8	262	1	FLGG_AGRFS	Q4438 agrobacteri
38	60	10.8	275	1	FLGG_AGRFS	Q90055 junonia coe
39	59.5	10.7	549	1	LIP2_CANRU	P32947 candida rug
40	59.5	10.7	718	1	CDG1_BACSS	P31747 bacillus sp
41	59.5	10.7	826	1	YEHF_ECOLI	P53341 escherichia
42	59.5	10.6	461	1	TBG_NEUCOR	P53377 neurospora
43	59	10.6	740	1	GNT5_RAT	Q08334 rattus norv
44	59	10.6	741	1	GNT5_RAT	Q08328 homo sapien
45	59	10.6	842	1	PHS2_RAT	P09812 rattus norv

ALIGNMENTS

RESULT 1
MPA2_CRYVA STANDARD; PRT, 514 AA.
AC P43212;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Possible polygalacturonase precursor (EC 3.2.1.15) (PG) (Pectinase)
DE (Major pollen allergen Cry j 2) (Cry j II)
OS Cryptomeria japonica (Japanese cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferales; Cupressaceae; Cryptomeria.
OX NCBI_TaxID=3369;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Pollen;
RX MEDLINE=95010777; PubMed=7926035;
RA Namba M., Kurose M., Toriye K., Hino K., Taniguchi Y., Fukuda S.,
RA Usui M., Kurimoto M.,
RT "Molecular cloning of the second major allergen, Cry j II, from
RT Japanese cedar pollen.";
RL PNAS Lett. 353:124-128 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pollen;
RX MEDLINE=94271186; PubMed=8002972;
RA Komiyama N., Sone T., Shimizu K., Morikubo K., Kino K.,
RT "cDNA cloning and expression of Cry j II the second major allergen of
RT Japanese cedar pollen.";
RL Biochem. Biophys. Res. Commun. 201:1021-1028 (1994).
RN [3]
RP SEQUENCE OF 55-64.
RX MEDLINE=90342988; PubMed=2382797;
RA Sakaguchi M., Inouye S., Tanai M., Ando S., Usui M., Matuhasi T.,
RT "Identification of the second major allergen of Japanese cedar
RT pollen.";
RL Allergy 45:309-312 (1990).
CC -!- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-
CC galactosiduronic linkages in pectate and other galacturonans.
CC -!- SUBCELLULAR LOCATION: SECRETED OR AMYLOPLAST (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
CC (POLYGALACTURONASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>)
CC or send an email to license@ebi.ac.uk.
CC -----
DR EMBL, D3765; BA007021.1; -;
DR EMBL, D29772; BA006172.1; -;
DR HSRP, P26509; IBBE.
DR InterPro, IPR000743; GR28.
DR Pfam, PF00295; Glyco_hydro_28; 1.

DR PROSITE: PS00502; POLYGALACTURONASE. 1.
 KW Hydrolyase; Glycosidase; Cell wall; Signal; Zymogen; Fruit ripening;
 KM Amyloplase; Glycoprotein; Allergen.
 FT SIGNAL 1 22
 FT PROPEP 23 45
 FT CHAIN 46 433
 FT PROPEP 514 434
 FT ACT SITE 278 278
 FT CARBOHYD 460 460
 FT CARBOHYD 472 472
 FT CARBOHYD 5 5
 FT CARBOHYD 12 12
 FT CARBOHYD 34 35
 FT CARBOHYD 37 37
 FT CARBOHYD 88 88
 FT CARBOHYD 98 98
 FT CARBOHYD 451 451
 FT CARBOHYD 454 454
 FT CARBOHYD 504 504
 FT CARBOHYD 507 507
 SO SEQUENCE 514 AA; 56645 MW; 624611C3FAD6302 CRC64;
 Query Match 20.7%; Score 115.5; DB 1; Length 514;
 Best Local Similarity 41.4%; Pred. No. 8.8e-06;
 Matches 29; Conservative 9; Mismatches 11; Indels 21; Gaps 4;
 Oy 12 PNRVPIKRVSNVTHGR-----RID--IFASKNPHLQKNTIGTGRWNRNRIWLOFA 62
 Db 96 PGNKFEV--VNNLFPNGCQHFTKVDGIILAYN-----PASWNNNNIWIQFA 143
 Oy 63 KLTGFTLMGR 72
 Db 144 KLTGFTLMGR 153
 RESULT 2
 SBP_CRYUA STANDARD; PRT; 374 AA.
 ID SBP_CRYUA
 AC P16632;
 DT 01-NOV-1990 (Rel. 16; Created)
 DT 01-NOV-1995 (Rel. 32; Last sequence update)
 DT 30-MAY-2000 (Rel. 39; Last annotation update)
 DE Sugil basic protein precursor (SBP) (Major allergen Cry j 1) (Cry j 1).
 OS Cryptomeria japonica (Japanese cedar).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.
 OC NCBI_TaxID=3369;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Pollen;
 RA MEDLINE=9418324; PubMed=8135802;
 RA Some T., Komiyama N., Shimizu K., Kusakabe T., Morikubo K.,
 RA Kino K.;
 RA "Cloning and sequencing of cDNA coding for Cry j I, a major allergen
 RT of Japanese cedar pollen.";
 RL Biochem. Biophys. Res. Commun. 199:619-625(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pollen;
 RA Namba M., Kurose M., Torigoe K., Fukuda S., Kurimoto M.;
 RA Submitted (JUL-1994) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 22-41.
 RC TISSUE=Pollen;
 RA MEDLINE=89031257; PubMed=3181436;
 RA Tanai M., Ando S., Ueki M., Kurimoto M., Sakaguchi M., Inouye S.,
 RA Matubasi T.;
 RA "N-terminal amino acid sequencing of a major allergen of Japanese cedar
 RT pollen (Cry j I).";
 RL FEBS Lett. 239:329-332(1988).
 RN [4]
 RP CARBOHYDRATES;
 RC TISSUE=Pollen;

RX MEDLINE=95003748; PubMed=7920021;
 RA Hijikata A., Matsunoto I., Kojima K., Ogawa H.;
 RT "Antigenicity of the oligosaccharide moiety of the Japanese cedar
 RL Int. Arch. Allergy Immunol. 105:198-202(1994).
 RN [5]
 RP STRUCTURE OF CARBOHYDRATES.
 RC TISSUE=Pollen;
 RA MEDLINE=95332249; PubMed=7608114;
 RA Hino K., Yamamoto S., Sano O., Taniguchi Y., Kohno K., Ueki M.,
 RA Fukuda S., Hanzawa H., Haruyama H., Kurimoto M.;
 RT "Carbohydrate structures of the glycoprotein allergen Cry j I from
 RL Japanese cedar (Cryptomeria japonica) pollen.";
 CC J. Biochem. 117:289-295(1995).
 CC -1- PTM: CONTAINS FUCOSE/XYLOSE--CONTAINING N-LINKED OLIGOSACCHARIDES.
 CC -1- DISEASE: THIS PROTEIN IS THE MAJOR ALLERGEN OF JAPANESE CEDAR
 CC POLLEN. THE MOST COMMON POLLEN ALLERGEN IN JAPAN.
 CC -1- MISCELLANEOUS: THE SEQUENCE OF CRY J I FORM A IS SHOWN HERE. FORM
 CC B DIFFERS IN SIX POSITIONS.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; D26544; BA05542.1; -;
 CC DR EMBL; D26545; BA05543.1; -;
 CC DR EMBL; D34639; BA07020.1; -;
 CC DR PIR; A44773; A44773.
 CC DR GLCOSUITE; P16632; -;
 CC DR InterPro; IPR002022; Amb allergen.
 CC DR Pfam; PF00544; pec lyase; 1.
 CC DR PRINTS; PR00807; AMBALLERGEN.
 CC DR Allergen; Glycoprotein; Multigene family; Signal.
 CC KW SIGNAL 1 21
 CC FT CHAIN 22 374
 CC FT CARBOHYD 158 158
 CC FT CARBOHYD 191 191
 CC FT CARBOHYD 293 293
 CC FT CARBOHYD 354 354
 CC FT CARBOHYD 12 12
 CC FT CARBOHYD 143 143
 CC FT CARBOHYD 202 202
 CC FT CARBOHYD 221 221
 CC FT CARBOHYD 358 358
 CC FT CARBOHYD 361 361
 CC FT CARBOHYD 374 AA; 40645 MW; 74AB2595024856F CRC64;
 SO SEQUENCE 374 AA; 40645 MW; 74AB2595024856F CRC64;
 Query Match 16.8%; Score 93.5; DB 1; Length 374;
 Best Local Similarity 51.4%; Pred. No. 0.0022;
 Matches 19; Conservative 6; Mismatches 7; Indels 5; Gaps 1;
 Oy 53 KNNRIWLOFAKLTGFTLMGRRLKMPYIAGYKTFDGR 89
 Db 85 RDRPLWITIFS-----GNMNIKLMKPMYIAGYKTFDGR 116
 RESULT 3
 MPAL_JUNAS STANDARD; PRT; 367 AA.
 ID MPAL_JUNAS
 AC P81294; O92NU7;
 DT 16-OCT-2001 (Rel. 40; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 15-JUN-2002 (Rel. 41; Last annotation update)
 DE Major pollen allergen Jun a 1 precursor.
 OS Juniperus ashei (Ozark white cedar).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
 OX NCBI_TaxId=13101;
 [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 31-38; 42-50; 58-80; 88-94;
 RP 117-124; 134-140; 160-164; 256-263 AND 322-325.
 RC TISSUE=Pollen;
 RX MEDLINE=99414163; PubMed=10482836;
 RA Midoro-Horiuchi T.M., Goldblum R.M., Kurosky A., Wood T.G.,
 RA Brooks E.G.,
 RT "Molecular cloning of mountain cedar (Juniperus ashei) pollen major
 RT allergen, Jun a 1,"
 RT J. Allergy Clin. Immunol. 104:613-617(1999).
 RN [2]
 RP SEQUENCE OF 22-50.
 RC TISSUE=Pollen;
 RX MEDLINE=99414162; PubMed=10482835;
 RA Midoro-Horiuchi T., Goldblum R.M., Kurosky A., Goetz D.W.,
 RA Brooks E.G.,
 RT "Isolation and characterization of the mountain cedar (Juniperus
 RT ashei) pollen major allergen, Jun a 1,"
 RT J. Allergy Clin. Immunol. 104:608-612(1999).
 CC -1- DISEASE: THIS PROTEIN CAUSES SEVERAL SEASONAL ALLERGIC RHINITIS
 CC IN NORTH AMERICA.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF106663; AAD03609.1;
 CC DR EMBL: AF106662; AAD03608.1;
 CC DR InterPro: IPR002022; Amb_allergen.
 CC DR Pfam: PF00544; pec_lyase; 1.
 CC DR PRINTS: PRO0807; AMBALLERGEN.
 CC KW Allergen; Glycoprotein; Signal.
 CC FT SIGNAL 21
 CC FT CHAIN 1 22 367 MAJOR POLLEN ALLERGEN JUN A 1.
 CC FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SQ SEQUENCE 367 AA; 39824 MW; FC9B81E675662E49 CRC64;
 CC
 CC Query Match 15.3%; Score 85.5; DB 1; Length 367;
 CC Best Local Similarity 40.0%; Pred. No. 0.018;
 CC Matches 16; Conservative 9; Mismatches 10; Indels 5; Gaps 1;
 CC
 CC QY 53 KNNRIMLOFAKLGTGRLMKPMYIAGYTFDGRVD 92
 CC Db 85 REKALWIFSQ-----NNNITKMKPLVYAGKHTIDGRAD 119
 CC
 CC RESULT 4
 CC MPAL_CUPAR STANDARD; PRT; 346 AA.
 CC ID MPAL_CUPAR
 CC AC Q9SGCG;
 CC DT 16-OCT-2001 (Rel. 40, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Major pollen allergen Cup a 1.
 CC OS Cupressus arizonica.
 CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 CC OX NCBI_TaxId=49011;
 CC [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=20571526; PubMed=11122214;
 CC RA Aceltuno E., Del Pozo V., Minguez A., Arieta I., Cortegano I.,
 CC Cardaba B., Gallardo S., Rojo M., Palomino P., Lahoz C.,

RT "Molecular cloning of major allergen from Cupressus arizonica pollen:
 RT Cup a 1,"
 RT Clin. Exp. Allergy 30:1750-1758(2000).
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AJ243570; CAB62551.1;
 CC DR InterPro: IPR002022; Amb_allergen.
 CC DR Pfam: PF00544; pec_lyase; 1.
 CC DR PRINTS: PRO0807; AMBALLERGEN.
 CC KW Allergen; Glycoprotein.
 CC FT CARBOHYD 127 127
 CC FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SQ SEQUENCE 346 AA; 37589 MW; F1281DCDAID5DFD0 CRC64;
 CC
 CC Query Match 14.8%; Score 82.5; DB 1; Length 346;
 CC Best Local Similarity 40.5%; Pred. No. 0.038;
 CC Matches 15; Conservative 9; Mismatches 8; Indels 5; Gaps 1;
 CC
 CC QY 53 KNNRIMLOFAKLGTGRLMKPMYIAGYTFDGR 89
 CC Db 64 REKALWIFSQ-----NNNITKMKPLVYAGKHTIDGR 95
 CC
 CC RESULT 5
 CC GLGA_SYNY3 STANDARD; PRT; 477 AA.
 CC ID GLGA_SYNY3
 CC AC P74521;
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Glycogen synthase (EC 2.4.1.21) (Starch [bacterial glycogen]
 CC de synthase).
 CC GN GLGA OR SLO0945.
 CC OS Synecocystis sp. (strain PCC 6803).
 CC OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
 CC OX NCBI_TaxId=1148;
 CC [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=97061201; PubMed=8905231;
 CC RA Kaneke T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 CC Miyajima N., Hirosewa M., Sugiyama M., Sasamoto S., Kimura T.,
 CC Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Nario K.,
 CC Okumura S., Shimo S., Takeuchi C., Wada T., Watanabe A.,
 CC Yamada M., Yasuda M., Tabata S.,
 CC RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synecocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions,"
 RT DNA Res. 3:109-136(1996).
 CC -1- FUNCTION: Synthesizes alpha-1,4-glucan chains using ADP-glucose.
 CC -1- CATALYTIC ACTIVITY: ADP-glucose + (1,4)-alpha-D-glucosyl (N) =
 CC ADP + (1,4)-alpha-D-glucosyl (N+1)
 CC -1- PATHWAY: Glycogen biosynthesis; second step.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; D90915; BAA18625.1; -
 DR InterPro; IPR001296; Glycos transf. 1.
 DR Pfam; PF00534; Glycos transf. 1; 1.
 KW Glycogen biosynthesis; Transferase; Glycosyltransferase;
 KM Complete proteome.
 FT BINDING 15 15 ADP-GLUCOSE (BY SIMILARITY).
 SQ SEQUENCE 477 AA; 54346 MW; 40C1B6013938ED32 CRC64;
 Query Match 13.5%; Score 75.5; DB 1; Length 477;
 Best Local Similarity 26.7%; Pred. No. 0.35;
 Matches 23; Conservative 15; Mismatches 29; Indels 19; Gaps 4;
 QY 8 NQGPENRERFTRKVSNTIIGRRIDIPASKNFHLOKNTIGRRKNNRIWLOFAKLTGR 67
 DB 31 HQLGHVAVFMPYVGYFI---GDKIDV-----PKRPV-----WKGEMFQFVAVQSY 74
 QY 68 TLMGRRLKMPYIAGYKTFPDRYDG 93
 DB 75 L--PDTKIPLYLFGHPAFDSRRITG 97
 RESULT 6
 MPAL CHAOS STANDARD; PRT; 375 AA.
 AC Q96385;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Major pollen allergen Cha o 1 precursor.
 OS Chamaecyparis obtusa (Japanese cypress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae;
 CC Chamaecyparis.
 CC NCBI_TaxID=13415;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Pollen.
 RX MEDLINE=96265194; PubMed=8676896;
 RA Suzuki M., Komiyama N., Itoh H., Sone T., Kuno K., Takagi I.,
 RA Ohta N.;
 RT "Purification, characterization and molecular cloning of Cha o 1, a
 RT major allergen of Chamaecyparis obtusa (Japanese cypress) pollen.";
 RL Mol. Immunol. 33:451-460(1996).
 CC -1- SIMILARITY: BELONGS TO THE POLYSCCHARIDE LYASE FAMILY 1.
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ldb-sib.ch/announce/>
 CC or send an email to license@ldb-sib.ch).
 CC
 CC EMBL; D45404; BAA08246.1; -
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec_lyase.1.
 DR PRINTS; PR00807; AMBALYERGEN.
 KW Allergen; Glycoprotein; Signal.
 FT SIGNAL 1 21
 FT CHAIN 1 21 MAJOR POLLEN ALLERGEN CHA O 1.
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 375 AA; 40258 MW; 81CD91DF7066DBBF CRC64;
 Query Match 13.4%; Score 74.5; DB 1; Length 375;
 Best Local Similarity 40.5%; Pred. No. 0.35;
 Matches 15; Conservative 7; Mismatches 10; Indels 5; Gaps 1;
 QY 53 KNNRIWLOFAKLTGFTLMGRRLKMPYIAGYKTFDGR 89

DB 85 RERSLWIFSK-----NNTIKLNPITYAGKTTIDGR 116
 RESULT 7
 DEEP SOLTU STANDARD; PRT; 576 AA.
 AC Q06803;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 4-alpha-glucanotransferase, chloroplast precursor (EC 2.4.1.25)
 DE (Amylomalase) (Disproportionating enzyme) (D-enzyme).
 GN DEEP.
 OS Solanum tuberosum (Potato).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 CC NCBI_TaxID=4113;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 53-57, 174-183 AND 247-259.
 RC STRAIN=cv. May Queen; TISSUE=tuber;
 RX MEDLINE=93123262; PubMed=7678257;
 RA Takeba T., Yamasaki M., Okada S., Smith S.M.;
 RT "Disproportionating enzyme (4-alpha-glucanotransferase; EC 2.4.1.25)
 RT of potato. Purification, molecular cloning, and potential role in
 RT starch metabolism.";
 RL J. Biol. Chem. 268:1391-1396(1993).
 CC -1- FUNCTION: MAY ACT DURING STARCH BREAKDOWN TO CONVERT SMALL
 CC OLIGOSACCHARIDES INTO LARGER MOLECULES UPON WHICH STARCH
 CC PHOSPHORYLASE CAN ACT, OR MAY CHANGE THE STRUCTURE OF STARCH
 CC MOLECULES AND GRAIN ARCHITECTURE BY MODIFYING CHAIN LENGTH, OR
 CC MAY GENERATE FROM STARCH AND GLUCOSE OLIGOSACCHARIDES WHICH
 CC CAN SERVE EITHER AS PRIMERS FOR NEW STARCH PHOSPHOENZYME.
 CC -1- CATALYTIC ACTIVITY: Transfers a segment of a (1,4)-alpha-D-glucan
 CC to a new 4-position in an acceptor, which may be glucose or (1,4)-
 CC alpha-D-glucan.
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST OR AMYLOPLAST.
 CC -1- TISSUE SPECIFICITY: PRESENT IN LEAVES, STEMS, ROOTS, AND STOLONS
 CC BUT IS MOST ABUNDANT IN DEVELOPING AND MATURE TUBERS.
 CC -1- SIMILARITY: BELONGS TO THE DISPROPORTIONATING ENZYME FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ldb-sib.ch/announce/>
 CC or send an email to license@ldb-sib.ch).
 CC
 CC EMBL; X68664; CAA48630.1; -
 DR PIR; A45049; A45049.
 DR HSSP; O87172; 1CWI.
 DR InterPro; IPR003385; 4A-glucanotrans.
 DR Pfam; PF02446; 4A-glucanotrans. 1.
 DR TIGRFAMs; TIGR00217; maiQ. 1.
 KW Transferase; Glycosyltransferase; Carbohydrate metabolism;
 KW Amyloplast; Chloroplast; Transit peptide.
 FT TRANSIT 1 52 CHLOROPLAST.
 FT CHAIN 53 576 4-ALPHA-GLUCANOTRANSFERASE.
 SQ SEQUENCE 576 AA; 64950 MW; A0D16F3A546307BB CRC64;
 Query Match 12.7%; Score 71; DB 1; Length 576;
 Best Local Similarity 35.1%; Pred. No. 1.4;
 Matches 20; Conservative 10; Mismatches 13; Indels 14; Gaps 4;
 QY 31 IDIFASKNFHLOKNTIGRRKNNRIWLOFAKLTGFTLMGRRLKMPYIAGYKTFD 87
 DB 262 IDIFIAQQLFQ-----RQMKVVR---DYARSKGISING---DMPITV--GYSHAD 304
 RESULT 8
 A29B_DROME

ID A298.DROME STANDARD; PRT; 234 AA.
 AC Q6197; Q9VT3; Q9TW05; Q9TW06; Q9TW07; Q9U976; Q9U977; Q9U978;
 AC Q9U979; Q9U905;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Accessory gland protein Acp29Ab precursor.
 GN ACP29AB OR CG17797.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pelegronata; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 CX NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
 RC STRAIN=Canton-S; TISSUE=Male accessory gland;
 RX MEDLINE=98135120; PubMed=9474779;
 RA Wolfner M.F., Harada H.A., Bertam M.J., Stelick T.J., Kraus K.W.,
 RA Kalb J.M., Lung Y.O., Neubaum D.M., Park M., Tzeng U.K.;
 RT "New genes for male accessory gland proteins in Drosophila
 RT melanogaster";
 RL Insect Biochem. Mol. Biol. 27:825-834(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Various strains;
 RX MEDLINE=9282496; PubMed=10353898;
 RA Agnede M.;
 RT "Positive selection drives the evolution of the Acp29Ab accessory
 RT gland protein in Drosophila";
 RL Genetics 152:543-551(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Various strains;
 RX MEDLINE=20556153; PubMed=1102381;
 RA Begun D.J., Whitely P., Todd B.L., Waldrip-Dail H.M., Clark A.G.;
 RT "Molecular population genetics of male accessory gland proteins in
 RT Drosophila";
 RL Genetics 156:1879-1888(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Gelink S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George K.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Chame M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Bolshakov S.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.B., Downes W., Dugan-Kocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz N.S., Gelbart W.M., Glasser K.,
 RA Foeller C., Gabrielian A.E., Garay N.S., Gerhart M., Fleischmann W.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Harris M.,
 RA Hoeltz D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jajati M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Liao P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Lin X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Moharir C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclik J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard U., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtk R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers R.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: RESPONSIBLE FOR PHYSIOLOGICAL AND BEHAVIORAL CHANGES IN
 CC MATED FEMALE FLIES.
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -1- TISSUE SPECIFICITY: MAIN CELLS OF THE ACCESSORY GLAND AND IN
 CC SEMINAL FLUID.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U85758; AAB96382.1; -
 CC EMBL; AJ240513; CAB53187.1; -
 CC EMBL; AJ240514; CAB53188.1; -
 CC EMBL; AJ240515; CAB53189.1; -
 CC EMBL; AJ240516; CAB53190.1; -
 CC EMBL; AJ240517; CAB53191.1; -
 CC EMBL; AJ240518; CAB53192.1; -
 CC EMBL; AJ240519; CAB53193.1; -
 CC EMBL; AJ240520; CAB53194.1; -
 CC EMBL; AJ240521; CAB53195.1; -
 CC EMBL; AJ240522; CAB53196.1; -
 CC EMBL; AJ240523; CAB53197.1; -
 CC EMBL; AJ240524; CAB53198.1; -
 CC EMBL; AJ240525; CAB53199.1; -
 CC EMBL; AJ240526; CAB53200.1; -
 CC EMBL; AJ240527; CAB53201.1; -
 CC EMBL; AJ240528; CAB53202.1; -
 CC EMBL; AJ240529; CAB53203.1; -
 CC EMBL; AJ240530; CAB53204.1; -
 CC EMBL; AJ240531; CAB53205.1; -
 CC EMBL; AJ240532; CAB53206.1; -
 CC EMBL; AJ240533; CAB53207.1; -
 CC EMBL; AJ240534; CAB53208.1; -
 CC EMBL; AJ240535; CAB53209.1; -
 CC EMBL; AJ240536; CAB53210.1; -
 CC EMBL; AJ240537; CAB53211.1; -
 CC EMBL; AJ240538; CAB53212.1; -
 CC EMBL; AJ240539; CAB53213.1; -
 CC EMBL; AJ240540; CAB53214.1; -
 CC EMBL; AJ240541; CAB53215.1; -
 CC EMBL; AJ240542; CAB53216.1; -
 CC EMBL; AJ240543; CAB53217.1; -
 CC EMBL; AJ240544; CAB53218.1; -
 CC EMBL; AJ240545; CAB53219.1; -
 CC EMBL; AJ240546; CAB53220.1; -
 CC EMBL; AJ240547; CAB53221.1; -
 CC EMBL; AJ240548; CAB53222.1; -
 CC EMBL; AJ240549; CAB53223.1; -
 CC EMBL; AJ240550; CAB53224.1; -
 CC EMBL; AJ240551; CAB53225.1; -
 CC EMBL; AY010527; AAG32585.1; -
 CC EMBL; AY010528; AAG32586.1; -
 CC EMBL; AY010529; AAG32587.1; -
 CC EMBL; AY010530; AAG32588.1; -
 CC EMBL; AY010531; AAG32589.1; -
 CC EMBL; AY010532; AAG32590.1; -
 CC EMBL; AY010533; AAG32591.1; -
 CC EMBL; AY010534; AAG32592.1; -
 CC EMBL; AY010535; AAG32593.1; -
 CC EMBL; AY010536; AAG32594.1; -
 CC EMBL; AY010537; AAG32595.1; -

```

Ra- Lettellell P., Louis E.J., Macri C., Mardis E., Meneses S., Mouser L.,
Rn Nman M., Rifkin L., Riles L., St Peter H., Teveskie E., Vaughan K.,
Ra Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
Rt Vaudin M.;
RL "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
viii.";
RT Science 265:2077-2082(1994).
CC -1- SIMILARITY: CONTAINS 2 BRCT DOMAINS.
-----
Cc This SWISS-PROT entry is copyright. It is produced through a collaboration
Cc between the Swiss Institute of Bioinformatics and the EMBL outstation -
Cc the European Bioinformatics Institute. There are no restrictions on its
Cc use by non-profit institutions as long as its content is in no way
Cc modified and this statement is not removed. Usage by and for commercial
Cc entities requires a license agreement (See http://www.isb.ch/announce/
Cc or send an email to license@isb-sib.ch).
Cc -----
Cc EMBL, U10397, AAB68978.1, -.
DR PIR, S46755, S46755.
DR SGD, S0001197, ESCA.
DR InterPro, IPR001357, BRCT.
DR Pfam, PF00533, BRCT, 5.
DR SMART, SMO0292, BRCT, 4.
DR PROSITE, PS50172, BRCT, 2.
DR Hypothetical protein; Repeat.
KW DOMAIN 117 210 BRCT 1.
FT DOMAIN 370 466 BRCT 2.
FT FT 767931285B52580 CRC64;
SQ SEQUENCE 1070 AA; 123017 MW; 767931285B52580 CRC64;
Query March 12.1%; Score 67.5; DB 1; Length 1070;
Best Local Similarity 27.4%; Pred. No. 7.3; Gaps 4;
Matches 20; Conservative 15; Mismatches 23; Indels 15;
Qy 2 KUTVAENGFGRNRVRIKVSNTIIIGRIDIPASKNFHL-QKNITG---TGRRWK--- 53
Db 379 ELTVAVTYVFSGRFYIQRLAVE-IIGSGSTPELTRKNTHLTSTIIGKKFVKAKWSLDP 437
Qy 54 -----NRRIMQL 60
Db 438 QNAIVTYHMMWLE 450
RESULT 10
Y12K SMSV4 STANDARD; FRT; 110 AA.
AC P36289;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Possible 12 kDa nucleic acid-binding protein.
OS San Miguel sea lion virus (serotype 4) (SMSV 4).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesivirus.
OX NCBI_TaxID=36407;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92410750; PubMed=1529644;
RA Neill J.D.;
RT "Nucleotide sequence of the capsid protein gene of two serotypes of
RT San Miguel sea lion virus: identification of conserved and non-
RT conserved amino acid sequences among calicivirus capsid proteins.";
RL Virus Res. 24:211-222(1992).
CC -1- SIMILARITY: TO FELINE CALCICIVIRUS 12 kDa PROTEIN.
-----
Cc This SWISS-PROT entry is copyright. It is produced through a collaboration
Cc between the Swiss Institute of Bioinformatics and the EMBL outstation -
Cc the European Bioinformatics Institute. There are no restrictions on its
Cc use by non-profit institutions as long as its content is in no way
Cc modified and this statement is not removed. Usage by and for commercial
Cc entities requires a license agreement (See http://www.isb-sib.ch/announce/
Cc or send an email to license@isb-sib.ch).
Cc -----
Cc EMBL, M67482; AAA16221.1; -.

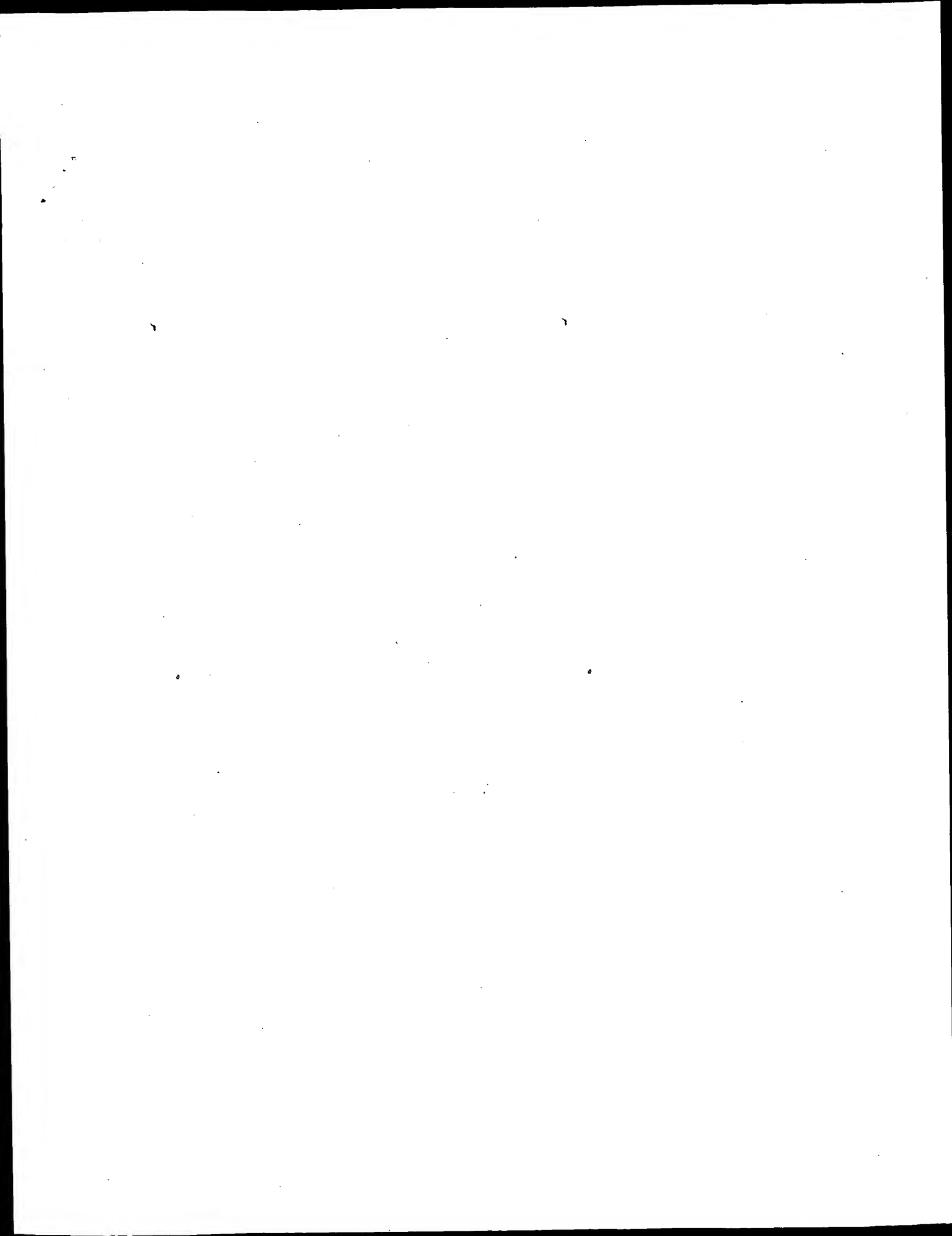
```

1

AC Q32131;
15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)

15-JUN-1998 (Rel. 36, Last annotation update)
 NMDH-plastoquinone oxidoreductase chain 5, chloroplast (EC 1.6.5.3)
 (fragment).
 DE NMDH.
 GN NDHF.
 OS Digitalis grandiflora (Yellow foxglove).
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Lamiales; Veronicaceae; Digitalis.
 NCBI_TaxID=38791;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Olmstead R.G., Reeves P.A.;
 RT "Evidence for the polyphyly of the Scrophulariaceae based on
 chloroplast rbcL and ndhF sequences." f
 Ann. Mo. Bot. Gard. 82:176-193 (1995).
 CC -1- CATALYTIC ACTIVITY: NMDH + plastoquinone = NMD(+) + plastoquinol.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L36399; AAA84203.1; -
 DR InterPro; IPR001750; Oxidored_q1.
 DR InterPro; IPR002128; Oxidored_q1_N.
 DR InterPro; IPR001516; Oxidored_q1_N.
 DR Pfam; PF00361; Oxidored_q1_1.
 DR Pfam; PF00662; Oxidored_q1_N; 1.
 DR Pfam; PF01010; Oxidored_q1_C; 1.
 KM Oxidoreductase; NAD; Plastoquinone; Chloroplast.
 FT NON TER 1 1
 FT NON TER 699 699
 SQ SEQUENCE 699 AA; 79309 MW; 5685FBA561C63D01 CRC64;
 Query Match 11.7%; Score 65.5; DB 1; Length 699;
 Best Local Similarity 30.3%; Pred. No. 7.8;
 Matches 23; Conservative 9; Mismatches 25; Indels 19; Gaps 3;
 QY 33 IPAS-KNFILOKNTIGTGR-----MKNNIIMLOFAKLTGFTLMGRRLKMP 78
 DB 618 IPASFENFDLINSFVKTGKRSRWDKILTLNWSHNPAYIDVFTSTGSIKLSOLT 677
 QY 79 VIAGYKTFDGRVDGI 94
 DB 678 HF-----FDYVIDGI 688
 RESULT 14
 AMR DROR STANDARD; PRT; 493 AA.
 AC 07015;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alpha-amyase-related protein precursor (EC 3.2.1.1).
 GN AMYREL.
 OS Drosophila orna (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7233;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA da Lage J.-L.;
 RT Submitted (JAN-2000) to the EMBL/GenBank/DBD databases.
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
 CC linkages in oligosaccharides and polysaccharides.
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO

KNOWN AS THE ALPHA-AMYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U96158; AAC39108.2; -
 DR HSSP; P56634; 1JAE.
 DR Flybase; FBgn021266; Dore\Amyrel.
 DR InterPro; IPR000461; Alpha_amyase.
 DR Pfam; PF00128; alpha-amyase; 1.
 DR Pfam; PF02806; alpha-amyase_C; 1.
 DR PRINTS; PR00110; ALPHAMYLASE.
 KM Hydrolase; Glycosidase; Carbohydrate metabolism; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 493
 FT ACT_SITE 207 207
 FT ACT_SITE 211 211
 FT ACT_SITE 309 309
 FT ACT_SITE 47 103
 FT DISULFID 156 170
 FT DISULFID 417 440
 FT DISULFID 447 459
 FT DISULFID 493 5384
 SQ SEQUENCE 493 AA; 55384 MW; 87BE4183B3F1580 CRC64;
 Query Match 11.6%; Score 64.5; DB 1; Length 493;
 Best Local Similarity 28.7%; Pred. No. 6.9;
 Matches 31; Conservative 10; Mismatches 42; Indels 25; Gaps 6;
 QY 10 FGNRRVFIRKSVNVIHGRI-----DIFASKNFILOKNTIGTGRMKNNIIMLO-F 61
 DB 233 FPNHSRPF--FQEVLDHGHETVSRSDEYKDLGAVTFEPRFSE-IGNAFRGNALKMLQSW 289
 QY 62 AKLTGFTLMGRRLKMPVIAGYKTF-----DGRRVGIIAAYONPASWK 105
 DB 290 GTGCGFLPBGQAL-----TFVDNHNDORDAGAVLSKSPRYK 327
 RESULT 15
 RS9 NEIMA STANDARD; PRT; 130 AA.
 ID RS9 NEIMA
 AC Q9UQ29;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 30S ribosomal protein S9.
 GN RPSI OR NMA0375 OR NMB2056.
 OS Neisseria meningitidis (serogroup A), and
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 NCBI_TaxID=65699, 491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
 RX MEDLINE=20222556; PubMed=10761919;
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 RA Davies R.M., Davis P., Devlin K., Feltham T., Hamlin N., Holroyd S.,
 RA Jørgen K., Leach S., Moule S., Mungall K., Quail M.A.,
 RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
 RA Whitehead S., Spratt B.G., Barrall B.G.,
 RT "Complete DNA sequence of a serogroup A strain of Neisseria
 meningitidis Z2491."
 RT Nature 404:502-506 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / Serogroup B;
 RX MEDLINE=20157555; PubMed=10710307;
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,



GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:56:55 ; Search time 88.6974 Seconds
(without alignments)
243.919 Million cell updates/sec

Title: US-09-142-524D-2

Perfect score: 538
Sequence: 1 MKVTVAFNQFGPNRRVPIKR.....FDGRVDGIITAYQNPASWK 105

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP_archaea:*
2: SP_bacteria:*
3: SP_fungi:*
4: SP_human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mmc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unclassified:*
15: SP_virus:*
16: SP_bacteriophage:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93.5	16.8	374	10	Q8RUR1
2	91.5	16.4	507	10	Q9FY19
3	86.5	15.5	367	10	Q93XU6
4	85.5	15.3	367	10	Q9LUT2
5	85.5	15.3	367	10	Q9LUT1
6	85.5	15.3	367	10	Q9MAS6
7	85.5	15.3	367	10	Q9MAS4
8	85.5	15.3	367	10	Q9MAS3
9	85.5	15.3	367	10	Q9MAS2
10	85.5	15.3	367	10	Q93X51
11	84.5	15.1	398	16	Q9WZL9
12	81.5	14.6	367	10	Q9MAS5
13	81	14.5	219	10	Q9SIO9
14	77	13.8	219	10	Q93777
15	74	13.3	221	10	Q9LWK1
16	73.5	13.2	644	16	Q8YEX3

17	72.5	13.0	440	10	Q949D3
18	71.5	12.8	645	16	Q985Q4
19	71.5	12.8	700	8	Q9THR5
20	68.5	12.3	390	16	Q9THB6
21	68	12.2	167	2	Q9KHV1
22	68	12.2	199	2	Q8VP01
23	68	12.2	221	10	Q9LFR1
24	68	12.2	576	10	Q9LV91
25	68	12.2	690	11	Q9ESJ8
26	67	12.0	572	8	Q9TII7
27	67	12.0	662	11	Q88694
28	67	12.0	689	11	Q91VD5
29	67	12.0	695	8	Q9TK54
30	66.5	11.9	493	5	Q9BH28
31	66.5	11.9	493	5	Q9BN06
32	66.5	11.9	686	8	P92329
33	66.5	11.9	746	8	Q9TLB7
34	66.5	11.9	747	8	Q9TLA7
35	66.5	11.9	748	8	Q94VZ8
36	66.5	11.9	777	17	Q93763
37	66.5	11.9	1429	5	Q960A1
38	66	11.8	188	17	Q9Y9P2
39	66	11.8	745	8	Q9TLC1
40	65.5	11.7	602	16	Q8U966
41	65.5	11.7	699	8	Q9TLF3
42	65.5	11.7	700	8	Q9THN4
43	65.5	11.7	750	16	Q9CNC8
44	65	11.6	280	2	Q521Q4
45	65	11.6	335	2	Q9AHF2

ALIGNMENTS

RESULT 1
Q8RUR1 PRELIMINARY; PRT; 374 AA.
ID Q8RUR1
AC Q8RUR1;
DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Cry j 1 precursor.
GN Cry j 1.1 OR Cry j 1.2.
OS Cryptomeria japonica (Japanese cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.
OX NCBI_TaxID=3369;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=POLLEN;
RA Futamura N., Shinohara K.;
RT "Isolation and characterization of cDNAs encoding major allergen Cry j
RT 1 from Cryptomeria japonica pollen."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB081309; BAB86286.1; -
DR EMBL; AB081310; BAB86287.1; -
KW Signal.
FT SIGNAL. 1 21 POTENTIAL.
FT CHAIN. 22 374 CRY j 1.
SQ SEQUENCE 374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;

Query Match 16.8%; Score 93.5; DB 10; Length 374;
Best Local Similarity 51.4%; Pred. No. 0.017; 7; Indels 5; Gaps 1;
Matches 19; Conservative 6; Mismatches 7; Indels 5; Gaps 1;

QY 53 KNNRIWLPQAKLTGFTLMGRLLKMPYIAGYTFDGR 89
DB 85 RDRPLWIRFS-----GNNWIKLKMMPYIAGYTFDGR 116

RESULT 2
Q9FY19 PRELIMINARY; PRT; 507 AA.

AC Q9FY19; 01-MAR-2001 (TREMBLREL. 16, Created)
 DT 01-MAR-2001 (TREMBLREL. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBLREL. 21, Last annotation update)
 DE Pollen major allergen 2 protein precursor.
 GN UN2.
 OS Juniperus ashei (Ozark white cedar).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
 RX NCBI_TaxID=1101;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MALE POLLEN;
 RX MEDLINE=20403896; PubMed=10944464;
 RA Yokoyama M., Miyahara M., Shimizu K., Kino K., Tsunoo H.;
 RT "Purification, identification and cDNA cloning of Jun a 2, the second
 RT major allergen of mountain cedar pollen."
 RL Biochem Biophys Res Commun 275:195-202(2000).
 CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
 CC (POLYGALACTURONASES).
 DR EMBL; AJ404653; CAC05582.1; -.
 DR HSSP; P26509; 1BHE.
 DR InterPro; IPR000743; GH28.
 DR InterPro; IPR000408; Reg chr condens.
 DR Pfam; PF00295; Glyco_Hydro_28; 1.
 DR PROSITE; PS00502; POLYGALACTURONASE; UNKNOWN_1.
 DR PROSITE; PS00626; RCO1_2; UNKNOWN_1.
 KM Cell wall; Glycosidase; Hydrolyase; Signal.
 FT SIGNAL 1 54 POTENTIAL.
 SQ SEQUENCE 507 AA; 55730 MW; 2B2E0AA5E5958FE5A CRC64;
 Query Match 16.4%; Score 91.5; DB 10; Length 507;
 Best Local Similarity 31.5%; Pred. No. 0.041;
 Matches 29; Conservative 14; Mismatches 26; Indels 23; Gaps 5;
 QY 12 ENRAREYKRVSVVTHGRIDIFASGNFLQNTIGT---GRMKNRIMWOPAKLTGFT 68
 DB 97 PANKFTV--VNDLVRYGPGCPHFSEK---VDGTTAAIPDPKMKSKIMHFAKLTDFN 150
 QY 69 LMGRLKMPMYIAGY-----KTFDGRV 91
 DB 151 LMGTV-----IDGQNRWMSDCKTINRTV 177
 RESULT 3
 Q93XL6 PRELIMINARY; PRT; 367 AA.
 AC Q93XL6; 01-DEC-2001 (TREMBLREL. 19, Created)
 DT 01-JUN-2002 (TREMBLREL. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLREL. 21, Last annotation update)
 DE Putative allergen Cup a 1 precursor.
 GN CUP A 1.
 OS Cupressus arizonica.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 RX NCBI_TaxID=49011;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=POLLEN;
 RA Buteront C., Di Felice G., Pini C.;
 RT "Cloning of Cupressus arizonica major allergen."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBS databases.
 DR EMBL; AJ278498; CAC37790.2; -.
 KW SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 367 PUTATIVE ALLERGEN CUP A 1.
 SQ SEQUENCE 367 AA; 39809 MW; AFE97260423A9F28 CRC64;
 Query Match 15.5%; Score 86.5; DB 10; Length 367;
 Best Local Similarity 40.0%; Pred. No. 0.11;
 Matches 16; Conservative 9; Mismatches 10; Indels 5; Gaps 1;

QY 53 KNNRIMLOFAKLTGFTLMGRRLKMPMYIAGYKTFDGRV 92
 DB 85 REKALWIIFSQ-----NMNIXKLMPLYVAGHKTIDGRAD 119
 RESULT 4
 Q9LTL2 PRELIMINARY; PRT; 367 AA.
 AC Q9LTL2; 01-OCT-2000 (TREMBLREL. 15, Created)
 DT 01-OCT-2000 (TREMBLREL. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLREL. 21, Last annotation update)
 DE Pollen major allergen 1.2.
 OS Juniperus virginiana (Eastern red cedar).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
 RX NCBI_TaxID=39584;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21315424; PubMed=11422137;
 RA Midoro-Horiuchi T.M., Goldblum R.M., Brooks E.G.;
 RT "Identification of mutations in the genes for the pollen allergens of
 RT eastern red cedar (Juniperus virginiana)."
 RL Clin. Exp. Allergy 31:771-778(2001).
 DR EMBL; AF151427; AAF80164.1; -.
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec lyase; 1.
 DR PRINTS; PRO0807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39708 MW; DCBD1981A74E4711 CRC64;
 Query Match 15.3%; Score 85.5; DB 10; Length 367;
 Best Local Similarity 40.0%; Pred. No. 0.14;
 Matches 16; Conservative 9; Mismatches 10; Indels 5; Gaps 1;
 QY 53 KNNRIMLOFAKLTGFTLMGRRLKMPMYIAGYKTFDGRV 92
 DB 85 REKALWIIFSQ-----NMNIXKLMPLYVAGHKTIDGRAD 119
 RESULT 5
 Q9LTL1 PRELIMINARY; PRT; 367 AA.
 AC Q9LTL1; 01-OCT-2000 (TREMBLREL. 15, Created)
 DT 01-OCT-2000 (TREMBLREL. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLREL. 21, Last annotation update)
 DE Pollen major allergen 1-1.
 OS Juniperus virginiana (Eastern red cedar).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
 RX NCBI_TaxID=39584;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21315424; PubMed=11422137;
 RA Midoro-Horiuchi T.M., Goldblum R.M., Brooks E.G.;
 RT "Identification of mutations in the genes for the pollen allergens of
 RT eastern red cedar (Juniperus virginiana)."
 RL Clin. Exp. Allergy 31:771-778(2001).
 DR EMBL; AF151429; AAF80166.1; -.
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec lyase; 1.
 DR PRINTS; PRO0807; AMBALLERGEN.
 SQ SEQUENCE 367 AA; 39708 MW; DCBD1981A74E4711 CRC64;
 Query Match 15.3%; Score 85.5; DB 10; Length 367;
 Best Local Similarity 40.0%; Pred. No. 0.14;
 Matches 16; Conservative 9; Mismatches 10; Indels 5; Gaps 1;

```

RESULT 6
QY 53 KNNRIWLOPAKLTGFTLMGRRLKMPYIAGYTFPDGRVD 92
DB 85 REKALWITFSQ-----NMNIKLMPLVYAGHKTIDGRAD 119

AC Q9M4S6; PRELIMINARY; PRT; 367 AA.
ID Q9M4S6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
GN CUP51.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX NCBI_TaxID=13469;
RN [1]
RP SEQUENCE FROM N.A.
RA Monsalve R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
RT of Cupressus sempervirens.";
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF257491; AAF72625.1; -.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39835 MW; B4E9C60108C2C5A3 CRC64;

Query Match 15.3%; Score 85.5; DB 10; Length 367;
Best Local Similarity 40.0%; Pred. No. 0.14;
Matches 16; Conservative 9; Mismatches 10; Indels 5; Gaps 1;

QY 53 KNNRIWLOPAKLTGFTLMGRRLKMPYIAGYTFPDGRVD 92
DB 85 REKALWITFSQ-----NMNIKLMPLVYAGHKTIDGRAD 119

AC Q9M4S4; PRELIMINARY; PRT; 367 AA.
ID Q9M4S4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
GN CUP51.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX NCBI_TaxID=13469;
RN [1]
RP SEQUENCE FROM N.A.
RA Monsalve R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
RT of Cupressus sempervirens.";
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF257493; AAF72627.1; -.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39934 MW; 974D3011D74E3DE6 CRC64;

Query Match 15.3%; Score 85.5; DB 10; Length 367;
Best Local Similarity 40.0%; Pred. No. 0.14;
Matches 16; Conservative 9; Mismatches 10; Indels 5; Gaps 1;

QY 53 KNNRIWLOPAKLTGFTLMGRRLKMPYIAGYTFPDGRVD 92
DB 85 REKALWITFSQ-----NMNIKLMPLVYAGHKTIDGRAD 119

AC Q9M4S3; PRELIMINARY; PRT; 367 AA.
ID Q9M4S3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)

```

```

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
GN CUP51.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX NCBI_TaxID=13469;
RN [1]
RP SEQUENCE FROM N.A.
RA Monsalve R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
RT of Cupressus sempervirens.";
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF257494; AAF72628.1; -.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39832 MW; B5DFBF5A61C07A53 CRC64;

Query Match 15.3%; Score 85.5; DB 10; Length 367;
Best Local Similarity 40.0%; Pred. No. 0.14;
Matches 16; Conservative 9; Mismatches 10; Indels 5; Gaps 1;

QY 53 KNNRIWLOPAKLTGFTLMGRRLKMPYIAGYTFPDGRVD 92
DB 85 REKALWITFSQ-----NMNIKLMPLVYAGHKTIDGRAD 119

AC Q9M4S2; PRELIMINARY; PRT; 367 AA.
ID Q9M4S2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
GN CUP51.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX NCBI_TaxID=13469;
RN [1]
RP SEQUENCE FROM N.A.
RA Monsalve R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
RT of Cupressus sempervirens.";
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF257495; AAF72629.1; -.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39819 MW; AE7E055A61C07A53 CRC64;

Query Match 15.3%; Score 85.5; DB 10; Length 367;
Best Local Similarity 40.0%; Pred. No. 0.14;
Matches 16; Conservative 9; Mismatches 10; Indels 5; Gaps 1;

QY 53 KNNRIWLOPAKLTGFTLMGRRLKMPYIAGYTFPDGRVD 92
DB 85 REKALWITFSQ-----NMNIKLMPLVYAGHKTIDGRAD 119

AC Q93X51; PRELIMINARY; PRT; 367 AA.
ID Q93X51;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Putative allergen Jun o 1.
GN JUN O 1.
OS Juniperus oxycedrus (Prickly juniper).

```

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
 OX NCBI_TaxId=69008;
 RN (1)
 RC SEQUENCE FROM N.A.
 RP TISSUE=POLLEN;
 RA Iacovacci P., Di Felice G., Pini C.;
 RT "Cloning of Juniperus oxycedrus major allergen";
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ293767; CAC48400.1; -
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 SQ SEQUENCE 367 AA; 39808 MW; 5D28204DBFD1B9D7 CRC64;

Query Match 15.3%; Score 85.5; DB 10; Length 367;
 Best Local Similarity 40.0%; Pred. No. 0.14;
 Matches 16; Conservative 9; Mismatches 10; Indels 5; Gaps 1;

DB 53 KNNRWLOPAKLTGFTLMGRRLKMPYIAGYKTFPGRRVD 92
 85 REKALMIIFSQ-----NMNKKLEMPLYVAGHKTIDRGAD 119

RESULT 11

O9WZL9 PRELIMINARY; PRT; 398 AA.
 AC O9WZL9;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE Lipopolysaccharide biosynthesis protein, putative.
 GN TM0760.
 OS Thermotoga maritima.
 OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
 OX NCBI_TaxId=2336;
 RN (1)
 RC SEQUENCE FROM N.A.
 RP STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gilm M.L., Dodson R.J.,
 RA Hafe D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eissen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 genome sequence of Thermotoga maritima";
 RL Nature 399:323-329 (1999).
 DR EMBL; AE001746; AAD35842.1; -
 DR TIGR; TM0760; -
 DR InterPro; IPR001296; Glycos_transf_1.
 DR Pfam; PF00534; Glycos_transf_1; 1.
 KW Complete proteome.
 SQ SEQUENCE 398 AA; 46448 MW; 28C5F84D55A25912 CRC64;

Query Match 15.1%; Score 84.5; DB 16; Length 398;
 Best Local Similarity 28.4%; Pred. No. 0.2;
 Matches 25; Conservative 11; Mismatches 35; Indels 17; Gaps 2;

OY 28 GRRIDIFASKNFHLQKNTIGTRAKNNRIW-----LQFAKLTGFTLMGRRLK 76
 DB 33 GHRVDIYICDFSHL-----TGKRWSETFGWSFKDGVDFIVERREXTGSLRLSSI 86
 OY 77 PMYIAGYKTFPGRRVDGIIAQNPAW 104
 DB 87 DYRNGRKKTIQKRYDVIIASSPHPSW 114

RESULT 12

O9M4S5 PRELIMINARY; PRT; 367 AA.
 AC O9M4S5;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Cup s 1 pollen allergen.
 GN CUPSL.
 OS Cupressus sempervirens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OX NCBI_TaxId=13469;
 RN (1)
 RC SEQUENCE FROM N.A.
 RP Monsalve R.I., Villalba M., Rodriguez R.;
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen
 of Cupressus sempervirens";
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF257492; AAF72626.1; -
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALALRGEN.
 SQ SEQUENCE 367 AA; 39894 MW; 5D56FC0E3263B741 CRC64;

Query Match 14.6%; Score 81.5; DB 10; Length 367;
 Best Local Similarity 37.5%; Pred. No. 0.4;
 Matches 15; Conservative 10; Mismatches 10; Indels 5; Gaps 1;

OY 53 KNNRWLOPAKLTGFTLMGRRLKMPYIAGYKTFPGRRVD 92
 85 REKALMIIFSQ-----NMNKKLEMPLYVAGHKTIDRGAD 119

RESULT 13

O9SIO9 PRELIMINARY; PRT; 219 AA.
 AC O9SIO9;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Putative synapcobrevin.
 GN AT2G25340.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxId=3702;
 RN (1)
 RC SEQUENCE FROM N.A.
 RP STRAIN=CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.U., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Vanden S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carreira A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Coppenhaver G.P., Preuss D., Niernan W.C., White O., Eissen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana";
 RL Nature 402:761-768 (1999).
 DR EMBL; AC007070; AAD3657.1; -
 DR InterPro; IPR001388; Synapcobrevin.
 DR Pfam; PF00957; synapcobrevin; 1.
 DR PRINTS; PR00219; SYNAPCOBREVN.
 DR ProDom; PD001229; Synapcobrevin; 1.
 SQ SEQUENCE 219 AA; 24965 MW; 7A63C85A140913B9 CRC64;

Query Match 14.5%; Score 81; DB 10; Length 219;
 Best Local Similarity 29.6%; Pred. No. 0.24;
 Matches 16; Conservative 14; Mismatches 22; Indels 2; Gaps 1;

OY 14 RRVFIRKRVSNVTHGRRIDIFASKNFHLQKNTIGTRRMK--NNRIWLOFAKLT 65
 AC 09LWK1; PRELIMINARY; PRT; 221 AA.
 DB 138 RGVMIENIDKVLDRGERLELVDTKTNANQGNTRFRKQTRFRFNNTVWVRNCKLT 191

RESULT 14

OY 049377 PRELIMINARY; PRT; 219 AA.
 AC 049377;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1999 (TREMBlrel. 09, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DB Synapcobrevin-like protein (AT4G32150/FlON7_40).
 GN FlON7_40 OR AT VAMP7C OR AT4G32150.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Koeltter P., Hempel S., Entian K.-D., Hobeisel J.,
 RA Mewes H.W., Mayer K.F.X., Scheller C.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Nikoloff D.M., Somerville C.R.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Rose M., Hempel S., Entian K.-D., Mewes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
 RA Banch J., Bower L., Carninci P., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamuya A.,
 RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
 RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
 RA Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;
 RT "Arabidopsis cDNA clones."
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL021636; CAAL6574.1; -
 DR EMBL; AF025332; AAD01748.1; -
 DR EMBL; AL161580; CAB79933.1; -
 DR EMBL; AF439840; AAL27509.1; -
 DR InterPro; IPR001388; Synapcobrevin.
 DR Pfam; PF00957; synapcobrevin; 1.
 DR PRINTS; PR00219; SYNAPTOBREVN.
 DR ProDom; PD001229; Synapcobrevin; 1.
 SQ SEQUENCE 219 AA; 25039 MW; 54C3BGF5C7A3D39B CRC64;

Query Match 13.8%; Score 77; DB 10; Length 219;
 Best Local Similarity 31.0%; Pred. No. 0.7;
 Matches 22; Conservative 16; Mismatches 27; Indels 6; Gaps 3;

OY 14 RRVFIRKRVSNVTHGRRIDIFASKNFHLQKNTI---GTGRMKNNRIWLOFAKLTGFTLM 70
 DB 138 RGVMIENIDKVLDRGERLELVDTKTNANQGNTRFRKQTRFRFNNTVWVRNCKLT--VLL 194
 OY 71 GRLKMPMYIA 81
 DB 195 ILLLVITYIA 205

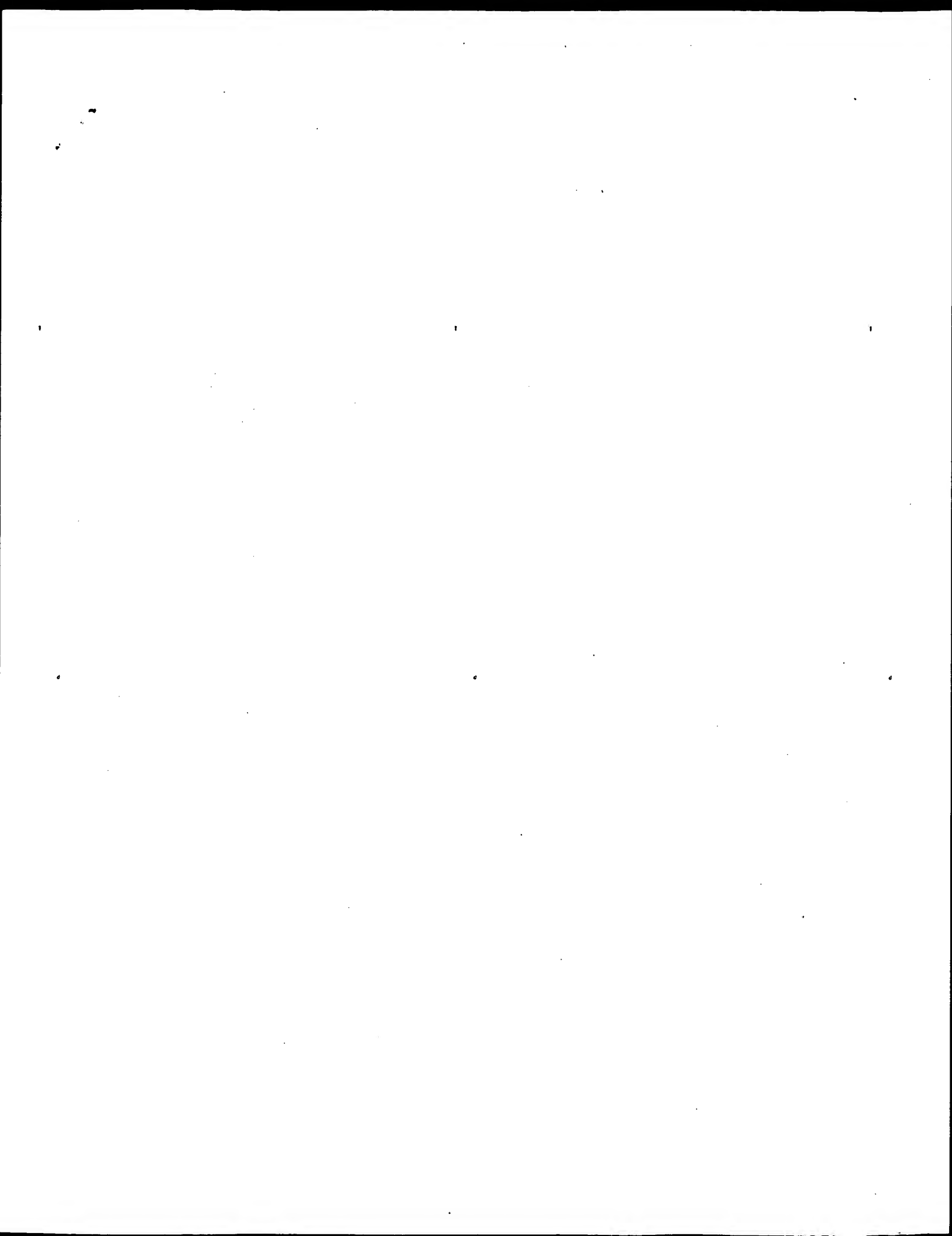
RESULT 15

ID 09LWK1 PRELIMINARY; PRT; 221 AA.
 AC 09LWK1;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE ESTs AU082579(S2069).
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzae; Oryza.
 OC NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
 RT clone: P0015E04."
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP002069; BAA95814.1; -
 DR InterPro; IPR001388; Synapcobrevin.
 DR Pfam; PF00957; synapcobrevin; 1.
 DR PRINTS; PR00219; SYNAPTOBREVN.
 DR ProDom; PD001229; Synapcobrevin; 1.
 SQ SEQUENCE 221 AA; 24970 MW; 1D573364B8E38571 CRC64;

Query Match 13.3%; Score 74; DB 10; Length 221;
 Best Local Similarity 29.1%; Pred. No. 1.6;
 Matches 23; Conservative 9; Mismatches 23; Indels 24; Gaps 3;

OY 14 RRVFIRKRVSNVTHGRRIDIFASKNFHLQKNTI---GTGRMKNNRIWLOFAKLT----- 65
 DB 139 RTVMINIDKVLDRGERLELVDTKTNANQGNTRFRKQTRFRFNNTVWVRNCKLTALIF 197
 OY 66 -----GFTL 69
 DB 198 LTVIIVVAVEMCHGFTL 216

Search completed: April 20, 2003, 13:12:57
 Job time : 91.6974 secs



GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using SW model

Run on: April 20, 2003, 12:54:05 ; Search time 127.382 Seconds
(without alignments)
-109.838 Million cell updates/sec

Title: US-09-142-524D-2

Perfect score: 558
Sequence: 1 MKVTVAFNQGFNRRVFIK.....FDGRVDGIIAAYNPASMK 105

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: A_Geneseq_101002:*
2: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
3: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
4: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
5: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
6: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
7: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
8: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
9: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
10: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
11: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
12: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
13: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
14: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
15: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
16: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
17: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
18: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
19: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
20: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
21: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
22: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
23: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
24: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	558	100.0	105	AAW27370	Multi-epitope pept
2	523.5	93.8	134	AAW27371	Multi-epitope pept
3	327.5	58.7	80	AAW27369	Multi-epitope pept
4	213	38.2	214	AAW27369	Multi-epitope pept
5	192.5	34.5	210	AAW27369	Multi-epitope pept
6	162	29.0	96	AAW27369	Multi-epitope pept
7	162	29.0	96	AAW27369	Multi-epitope pept
8	162	29.0	96	AAW27369	Multi-epitope pept
9	149.5	26.8	97	AAW27369	Multi-epitope pept
10	141.5	25.4	93	AAW27369	Multi-epitope pept

11	131.5	23.6	47	19	AAW80353	Sugi allergen prot
12	120.5	21.6	74	21	AAW23905	Artificial sequenc
13	120.5	21.6	94	21	AAW23906	Artificial sequenc
14	120.5	21.6	95	21	AAW23897	Artificial sequenc
15	120.5	21.6	95	21	AAW23901	Artificial sequenc
16	120.5	21.6	95	21	AAW23908	Artificial sequenc
17	120.5	21.6	99	21	AAW23902	Artificial sequenc
18	120.5	21.6	99	21	AAW23907	Artificial sequenc
19	120.5	21.6	99	21	AAW23909	Artificial sequenc
20	115.5	20.7	514	16	AAW74333	Cedar pollen aller
21	115.5	20.7	514	17	AAW81586	Japanese cedar pol
22	115.5	20.7	514	20	AAW25667	Cedar pollen aller
23	115	20.6	460	16	AAW69791	Japanese cedar all
24	115	20.6	514	15	AAW36901	Japanese cedar all
25	115	20.6	514	16	AAW69792	Japanese cedar pol
26	115	20.6	514	17	AAW36901	Japanese cedar pol
27	115	20.6	514	20	AAW25666	Japanese cedar all
28	106.5	19.1	71	21	AAW23878	Artificial sequenc
29	106.5	19.1	81	19	AAW80358	Artificial sequenc
30	106.5	19.1	81	21	AAW23874	Artificial sequenc
31	106.5	19.1	81	21	AAW23879	Artificial sequenc
32	106.5	19.1	81	21	AAW23895	Artificial sequenc
33	106.5	19.1	81	21	AAW23896	Artificial sequenc
34	106.5	19.1	81	22	AAW69094	Artificial sequenc
35	106.5	19.1	81	22	AAW69118	Cedar pollen aller
36	106.5	19.1	81	22	AAW69114	Cedar pollen aller
37	106.5	19.1	93	21	AAW23891	Artificial sequenc
38	106.5	19.1	93	21	AAW23896	Artificial sequenc
39	103	18.5	42	22	AAW84105	Cedar pollen aller
40	103	18.5	47	19	AAW80351	Immunomodulatory p
41	101.5	18.2	47	19	AAW80356	Sugi allergen prot
42	101	18.1	47	19	AAW80357	Sugi allergen prot
43	99	17.7	19	22	AAW84115	Amino acid sequenc
44	99	17.7	47	19	AAW80352	Sugi allergen prot
45	99	17.7	79	21	AAW23880	Artificial sequenc

ALIGNMENTS

RESULT 1	AAW27370	standard; peptide; 105 AA.
ID	AAW27370	
XX	AAW27370:	
AC	AAW27370:	
XX	AAW27370:	
DT	24-MAR-1998 (first entry)	
XX	Multi-epitope peptide used as immunotherapeutic agent #2.	
DE	Multi-epitope peptide; immunotherapeutic agent; allergic disease;	
XX	T-cell epitope region; allergen; lymphocyte; immunoglobulin E.	
KW	Synthetic.	
OS	WO9732600-A1.	
XX	12-SEP-1997.	
PD	10-MAR-1997; 97WO-UP00740.	
XX	10-MAR-1996; 96UP-0080702.	
PR	(MEIP) MEIJI MILK PROD CO LTD.	
XX	Dairiki K, Iwama A, Kino K, Kume A, Sone T;	
XX	WPI, 1997-470495/43.	
DR	Peptide immuno-therapeutic agent to treat allergic diseases -	
XX	contains multi-epitope peptide containing T cell epitope regions	
PT	from different allergens	
XX		

PS Claim 6; Page 31; 58bp; Japanese.

CC The present sequence represents a multi-epitope peptide which is used as
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 CC or more different allergens (preferably linked via arginine or lysine
 CC dimers), where the T cell epitope regions have a positivity index
 CC greater than 100 as measured in a patient group responding to the
 CC allergen; have at least 70% reactivity with lymphocytes from patients
 CC responding to the allergen; and are not reactive with immunoglobulin E
 CC (IgE) antibodies from patients responsive to the allergen. The agent can
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.

XX Sequence 105 AA;

Query Match 100.0%; Score 558; DB 18; Length 105;
 Best Local Similarity 100.0%; Pred. No. 3.1e-65;
 Matches 105; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKVTVAFNQFGNRRVFIKRVSNVLIHGRRIDIFASKNPHLQKNTIGTGRWKNNRIWLQ 60
 DB 1 MKVTVAFNQFGNRRVFIKRVSNVLIHGRRIDIFASKNPHLQKNTIGTGRWKNNRIWLQ 60

OY 61 FAKLTGFTLMGRRLKMPYIAGYKTFDGRVDTIIAAYQNPASMK 105
 DB 61 FAKLTGFTLMGRRLKMPYIAGYKTFDGRVDTIIAAYQNPASMK 105

RESULT 2

AAW27371
 ID AAW27371 standard; peptide; 134 AA.

AC AAW27371;

DT 24-MAR-1998 (first entry)

DE Multi-epitope peptide used as immunotherapeutic agent #3.

KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;
 KM T-cell epitope region; allergen; lymphocyte; immunoglobulin E.

OS Synthetic.

PN WO9732600-A1.

PD 12-SEP-1997.

PF 10-MAR-1997; 97WO-JP00740.

PR 10-MAR-1996; 96JP-0080702.

PA (MEIP) MEIJI MILK PROD CO LTD.

PI Dai-ichi K, Iwama A, Kuno K, Kume A, Sone T;

DR WPI; 1997-470495/43.

PT Peptide immuno:therapeutic agent to treat allergic diseases -
 PT contains multi-epitope peptide containing T cell epitope regions
 PT from different allergens

PS Claim 6; Page 32; 58bp; Japanese.

CC The present sequence represents a multi-epitope peptide which is used as
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 CC or more different allergens (preferably linked via arginine or lysine
 CC dimers), where the T cell epitope regions have a positivity index
 CC greater than 100 as measured in a patient group responding to the
 CC allergen; have at least 70% reactivity with lymphocytes from patients
 CC responding to the allergen; and are not reactive with immunoglobulin E
 CC (IgE) antibodies from patients responsive to the allergen. The agent can
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.

XX Sequence 134 AA;

Query Match 93.8%; Score 523.5; DB 18; Length 134;
 Best Local Similarity 78.4%; Pred. No. 1.4e-60;
 Matches 105; Conservative 0; Mismatches 0; Indels 29; Gaps 2;

OY 1 MKVTVAFNQFGNRRVFIKRVSNVLIHGRRIDIFASKNPHLQKNTIGTGRWKNNRIWLQ 60
 DB 1 MKVTVAFNQFGNRRVFIKRVSNVLIHGRRIDIFASKNPHLQKNTIGTGRWKNNRIWLQ 60

OY 61 FAKLTGFTLMGRRLKMPYIAGYKTFDGRVDTIIAAYQNPASMK 105
 DB 61 FAKLTGFTLMGRRLKMPYIAGYKTFDGRVDTIIAAYQNPASMK 105

OY 92 DGIHAYQNPASMK 105
 DB 121 DGIHAYQNPASMK 134

RESULT 3

AAW27369
 ID AAW27369 standard; peptide; 80 AA.

AC AAW27369;

DT 24-MAR-1998 (first entry)

DE Multi-epitope peptide used as immunotherapeutic agent #1.

KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;
 KM T-cell epitope region; allergen; lymphocyte; immunoglobulin E.

OS Synthetic.

PN WO9732600-A1.

PD 12-SEP-1997.

PF 10-MAR-1997; 97WO-JP00740.

PR 10-MAR-1996; 96JP-0080702.

PA (MEIP) MEIJI MILK PROD CO LTD.

PI Dai-ichi K, Iwama A, Kuno K, Kume A, Sone T;

DR WPI; 1997-470495/43.

PT Peptide immuno:therapeutic agent to treat allergic diseases -
 PT contains multi-epitope peptide containing T cell epitope regions
 PT from different allergens

PS Claim 6; Page 31; 58bp; Japanese.

CC The present sequence represents a multi-epitope peptide which is used as
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 CC or more different allergens (preferably linked via arginine or lysine
 CC dimers), where the T cell epitope regions have a positivity index
 CC greater than 100 as measured in a patient group responding to the
 CC allergen; have at least 70% reactivity with lymphocytes from patients
 CC responding to the allergen; and are not reactive with immunoglobulin E
 CC (IgE) antibodies from patients responsive to the allergen. The agent can
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.

XX Sequence 80 AA;

Query Match 58.7%; Score 327.5; DB 18; Length 80;
 Best Local Similarity 67.6%; Pred. No. 3.8e-35;
 Matches 71; Conservative 1; Mismatches 8; Indels 25; Gaps 2;

OY 1 MKVTVAFNQFGNRRVFIKRVSNVLIHGRRIDIFASKNPHLQKNTIGTGRWKNNRIWLQ 60

Db 1 MKVTVAFNQGPNNRRVFIKRVSNVLIHGRRIDIFASKNFHLQKNTIGTGRIS----- 53
 QY 61 FAKLTGFTLMGRRLKMPMYIAGYKTPGRVDDGIIAAYONPASWK 105
 Db 54 -LKLTSCKIA-----SRKVDGIIAAYONPASWK 80

RESULT 4

AAB69120
 ID AAB69120 standard; Protein; 214 AA.

AC AAB69120;
 XX

DT 23-APR-2001 (first entry)
 XX

DE Cedar pollen allergen T cell epitope derived protein SEQ ID NO:61.
 XX

KM Japanese cedar; Cryptomeria japonica; cedar pollen allergen;
 XX

OS Cryptomeria japonica.
 XX

PN JP2000327699-A.
 XX

PD 28-NOV-2000.
 XX

PF 15-MAR-2000; 2000JP-0071710.
 XX

PR 15-MAR-1999; 99JP-0068316.
 XX

PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 XX

DR WPI, 2001-185061/19.
 XX

DR N-PSDB; AAF59044.
 XX

PT Novel peptide and its use -
 XX

PS Example 11; Page 58-59; 75PP; Japanese.
 XX

CC The present invention describes a peptide, its complex, derivative or
 CC its polymerizate, where the peptide (I) has a formula of:
 CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7

CC where alpha1 to alpha7 = amino acid sequence selected from amino acid
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes
 CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen
 CC allergens. The peptide can be used in an antiasthma agent.
 CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used
 CC in the exemplification of the present invention.
 CC

XX Sequence 214 AA;
 SQ

Query Match 38.2%; Score 213; DB 22; Length 214;
 Best Local Similarity 41.0%; Pred. No. 1.6e-13;

Matches 55; Conservative 7; Mismatches 8; Indels 64; Gaps 5;

QY 1 MKVTVAFNQGPNNRRVFIKRVSNVLIHGRRIDIFASKNFHLQKNTIGTGR----- 50
 Db 30 MKVTVAFNQGP-----DIFASKNFHLQKNTIGTGR----- 70
 QY 51 -----RMKNKRIWL-----QPAKLTGFTLMGRRLKMPMYIAGYKTPGRV 91
 Db 71 HVANNNDPSGKTEGNGITTKKAENVEQFAKLTGFTLMGRA-----DPR-- 115
 QY 92 DGIIAAYONPASWK 105
 Db 116 -GIIAAYONPASWK 128

RESULT 5

AAB69103

ID AAB69103 standard; Protein; 210 AA.

AC AAB69103;
 XX

DT 23-APR-2001 (first entry)
 XX

DE Cedar pollen allergen T cell epitope derived protein SEQ ID NO:12.
 XX

KM Japanese cedar; Cryptomeria japonica; cedar pollen allergen;
 XX

OS Cryptomeria japonica.
 XX

PN JP2000327699-A.
 XX

PD 28-NOV-2000.
 XX

PF 15-MAR-2000; 2000JP-0071710.
 XX

PR 15-MAR-1999; 99JP-0068316.
 XX

PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 XX

DR WPI, 2001-185061/19.
 XX

DR N-PSDB; AAF59012.
 XX

PT Novel peptide and its use -
 XX

PS Claim 7; Page 39-40; 75PP; Japanese.
 XX

CC The present invention describes a peptide, its complex, derivative or
 CC its polymerizate, where the peptide (I) has a formula of:
 CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7
 CC where alpha1 to alpha7 = amino acid sequence selected from amino acid
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes
 CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen
 CC allergens. The peptide can be used in an antiasthma agent.
 CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used
 CC in the exemplification of the present invention.
 CC

XX Sequence 210 AA;
 SQ

Query Match 34.5%; Score 192.5; DB 22; Length 210;
 Best Local Similarity 39.6%; Pred. No. 7.5e-17;

Matches 53; Conservative 7; Mismatches 7; Indels 67; Gaps 6;

QY 1 MKVTVAFNQGPNNRRVFIKRVSNVLIHGRRIDIFASKNFHLQKNTIGTGR----- 50
 Db 29 MKVTVAFNQGP-----FASKNFHLQKNTIGTGR----- 67
 QY 51 -----RMKNKRIWL-----QPAKLTGFTLMGRRLKMPMYIAGYKTPGRV 90
 Db 68 HVANNNDPSGKTEGNGITTKKAENVEQFAKLTGFTLMGRA-----DPR-- 113
 QY 91 VDGIIAAYONPASWK 104
 Db 114 -GIIAAYONPASWK 125

RESULT 6

AAB69104

ID AAB69104 standard; Protein; 96 AA.

AC AAB69104;
 XX

DT 23-APR-2001 (first entry)
 XX

DE Cedar pollen allergen T cell epitope derived protein SEQ ID NO:13.
 XX

KM Japanese cedar; Cryptomeria japonica; cedar pollen allergen;
 XX

T cell epitope; antiasthma agent.

[illegible]

XX	RESULT 8
XX	AB69105
ID	AB69105 standard; Protein; 97 AA.
XX	
AC	AB69105;
XX	
DT	23-APR-2001 (first entry)
XX	
DE	Cedar pollen allergen T cell epitope derived protein SEQ ID NO:15.
XX	
KW	Japanese cedar; Cryptomeria japonica; cedar pollen allergen;
XX	T cell epitope; antiSugiPollinosis.
XX	
OS	Cryptomeria japonica.
OS	Synthetic.
XX	
PN	JP2000327699-A.
XX	
PD	28-NOV-2000.
XX	
PF	15-MAR-2000; 2000JP-0071710.
XX	
PR	15-MAR-1999; 99JP-0068316.
XX	
PA	(HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX	(SANY) SANKYO CO LTD.
DR	WPI: 2001-185061/19.
XX	
DR	N-PSDB; AAF59013.
XX	
PT	Novel peptide and its use -
XX	
PS	Claim 9; Page 41-42; 75pp; Japanese.
XX	
CC*	The present invention describes a peptide, its complex, derivative or
CC	its polymerizate, where the peptide (I) has a formula of:
CC	alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7
CC	where alpha1 to alpha7 = amino acid sequence selected from amino acid
CC	sequences ranging from 11-19 amino acids derived from T cell epitopes
CC	derived from cedar (Japanese cedar - Cryptomeria japonica) pollen
CC	allergens. The peptide can be used in an antiSugiPollinosis agent.

KW T cell epitope: sugi allergen proteins Cryj1; Cryj2; treatment;
 KM sugi-pollinosis; allergic reaction; pollen.
 XX Synthetic.
 OS
 XX JP10259198-A.
 PN
 XX 29-SEP-1998.
 PD
 XX 22-DEC-1997; 97JP-0353448.
 PF
 XX 24-DEC-1996; 96JP-0343441.
 PR
 XX (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA (SANY) SANKYO CO LTD.
 XX WPI; 1998-577037/49.
 DR
 XX A linked T cell epitope peptide - used for the treatment of
 PT sugi-pollinosis
 PS Claim 10; Page 5; 21pp; Japanese.
 XX
 CC AAW80339-58 represent epitopes for T cells, derived from the sugi
 CC allergen proteins Cryj1 (AAW80339-44, AAW80350-53 and AAW80356-58) and
 CC Cryj2 (AAW80345-49 and AAW80354-55). The peptides are useful for the
 CC treatment of sugi-pollinosis, an allergic reaction of the body to
 CC pollen.
 CC
 XX Sequence 47 AA;
 SQ
 Query Match 23.6%; Score 131.5; DB 19; Length 47;
 Best Local Similarity 33.7%; Pred. No. 1e-09;
 Matches 35; Conservative 0; Mismatches 0; Indels 69; Gaps 2;
 OY 1 MKVTAFNFGFNRRVFIKRVSNVLIHGRIDIFASKNFHLOKNTIGGRMKNNRWLQ 60
 DB 13 MKVTAFNFGF-----FASKNFHLOKNT----- 35
 OY 61 FKLNGFTLMGRRLKMPVYIAGYKTFDGRRDVGIIAAYONPASM 104
 DB 36 -----GIIAAYONPASM 47
 RESULT 12
 AAB23905
 ID AAB23905 standard; peptide; 74 AA.
 AC AAB23905;
 XX
 DT 17-JAN-2001 (first entry)
 DE Artificial sequence designed peptide #26.
 XX
 KM Peptide synthesis; chemical synthesis; solid phase synthesis.
 OS Synthetic.
 XX WO200055182-A1.
 PN
 XX 21-SEP-2000.
 PD
 XX 15-MAR-2000; 2000WO-JP01584.
 PF
 XX 15-MAR-1999; 99JP-0067917.
 PR (SANY) SANKYO CO LTD.
 PA
 XX Kawaguchi J, Serizawa N;
 PI WPI; 2000-602106/57.
 DR
 XX Chemically synthesizing a peptide by solid phase synthesis by adding

PT protective group to inactivate part of alpha amino group of an amino
 PT acid group -
 XX
 XX Example 6; Page 23; 38pp; Japanese.
 CC The present invention describes a method for chemically synthesizing a
 CC peptide by solid phase synthesis. The method comprises processing a
 CC carrier resin to which the C-terminal amino acid derivative of the
 CC target peptide has been preliminarily bonded is characterised by adding
 CC a protective group to inactivate a part of alpha-amino groups of an
 CC amino acid capable of undergoing a peptide extension reaction on the
 CC carrier resin. The method is useful for synthesizing a peptide by solid
 CC phase synthesis. Long chain peptides can be readily synthesised in with
 CC a high success ratio without using the segment condensation method or
 CC gene manipulations. The present sequence represents an amino acid
 CC peptide sequence given in an example from the present invention.
 SQ Sequence 74 AA;
 Query Match 21.6%; Score 120.5; DB 21; Length 74;
 Best Local Similarity 52.0%; Pred. No. 5.1e-08;
 Matches 26; Conservative 3; Mismatches 2; Indels 19; Gaps 1;
 OY 1 MKVTAFNFGFNRRVFIKRVSNVLIHGRIDIFASKNFHLOKNTIGGR 50
 DB 6 MKVTAFNFGF-----DIFASKNFHLOKNTIGGR 36
 RESULT 13
 AAB23906
 ID AAB23906 standard; peptide; 94 AA.
 AC AAB23906;
 XX
 DT 17-JAN-2001 (first entry)
 DE Artificial sequence designed peptide #27.
 XX
 KM Peptide synthesis; chemical synthesis; solid phase synthesis.
 OS Synthetic.
 XX WO200055182-A1.
 PN
 XX 21-SEP-2000.
 PD
 XX 15-MAR-2000; 2000WO-JP01584.
 PF
 XX 15-MAR-1999; 99JP-0067917.
 PR (SANY) SANKYO CO LTD.
 PA
 XX Kawaguchi J, Serizawa N;
 PI WPI; 2000-602106/57.
 DR
 XX Chemically synthesizing a peptide by solid phase synthesis by adding
 PT protective group to inactivate part of alpha amino group of an amino
 PT acid group -
 XX
 XX Example 6; Page 24; 38pp; Japanese.
 CC The present invention describes a method for chemically synthesizing a
 CC peptide by solid phase synthesis. The method comprises processing a
 CC carrier resin to which the C-terminal amino acid derivative of the
 CC target peptide has been preliminarily bonded is characterised by adding
 CC a protective group to inactivate a part of alpha-amino groups of an
 CC amino acid capable of undergoing a peptide extension reaction on the
 CC carrier resin. The method is useful for synthesizing a peptide by solid
 CC phase synthesis. Long chain peptides can be readily synthesised in with
 CC a high success ratio without using the segment condensation method or
 CC gene manipulations. The present sequence represents an amino acid
 CC peptide sequence given in an example from the present invention.

XX Sequence 94 AA;
SQ

Query Match 21.6%; Score 120.5; DB 21; Length 94;
Best Local Similarity 52.0%; Pred. No. 7.2e-08;
Matches 26; Conservative 3; Mismatches 2; Indels 19; Gaps 1;

QY 1 MKVTVAENQFGPNRRVFIKRVSNVTHGRRIDIFASKNFHLQKNTLTGTR 50
DB 26 MKVTVAENQFGP-----DIFASKNFHLQKNTLTGSK 56

RESULT 14

AAB23897 standard; peptide; 95 AA.

AC AAB23897;

DT 17-JAN-2001 (first entry)

DE Artificial sequence designed peptide SEQ ID NO:5.

KW Peptide synthesis; chemical synthesis; solid phase synthesis.

OS Synthetic.

PN WO200055182-A1.

PD 21-SEP-2000.

PF 15-MAR-2000; 2000WO-JP01584.

PR 15-MAR-1999; 99JP-0067917.

PA (SANY) SANKYO CO LTD.

PI Kawaguchi J, Serizawa N;

DR WPI; 2000-602106/57.

PT Chemically synthesizing a peptide by solid phase synthesis by adding
PT protective group to inactivate part of alpha amino group of an amino
PT acid group

PS Example 5; Page 34-35; 38pp; Japanese.

CC The present invention describes a method for chemically synthesizing a
CC peptide by solid phase synthesis. The method comprises processing a
CC carrier resin to which the C-terminal amino acid derivative of the
CC target peptide has been preliminarily bonded is characterized by adding
CC a protective group to inactivate a part of alpha-amino groups of an
CC amino acid capable of undergoing a peptide extension reaction on the
CC carrier resin. The method is useful for synthesizing a peptide by solid
CC phase synthesis. Long chain peptides can be readily synthesised in with
CC a high success ratio without using the segment condensation method or
CC gene manipulations. The present sequence represents an amino acid
CC peptide sequence given in an example from the present invention.

SQ Sequence 95 AA;

Query Match 21.6%; Score 120.5; DB 21; Length 95;
Best Local Similarity 52.0%; Pred. No. 7.2e-08;
Matches 26; Conservative 3; Mismatches 2; Indels 19; Gaps 1;

QY 1 MKVTVAENQFGPNRRVFIKRVSNVTHGRRIDIFASKNFHLQKNTLTGTR 50
DB 27 MKVTVAENQFGP-----DIFASKNFHLQKNTLTGSK 57

RESULT 15
AAB23901
ID AAB23901 standard; peptide; 95 AA.
XX

AC AAB23901;

DT 17-JAN-2001 (first entry)

DE Artificial sequence designed peptide #23.

KW Peptide synthesis; chemical synthesis; solid phase synthesis.

OS Synthetic.

PN WO200055182-A1.

PD 21-SEP-2000.

PF 15-MAR-2000; 2000WO-JP01584.

PR 15-MAR-1999; 99JP-0067917.

PA (SANY) SANKYO CO LTD.

PI Kawaguchi J, Serizawa N;

DR WPI; 2000-602106/57.

PT Chemically synthesizing a peptide by solid phase synthesis by adding
PT protective group to inactivate part of alpha amino group of an amino
PT acid group

PS Example 5; Page 21-22; 38pp; Japanese.

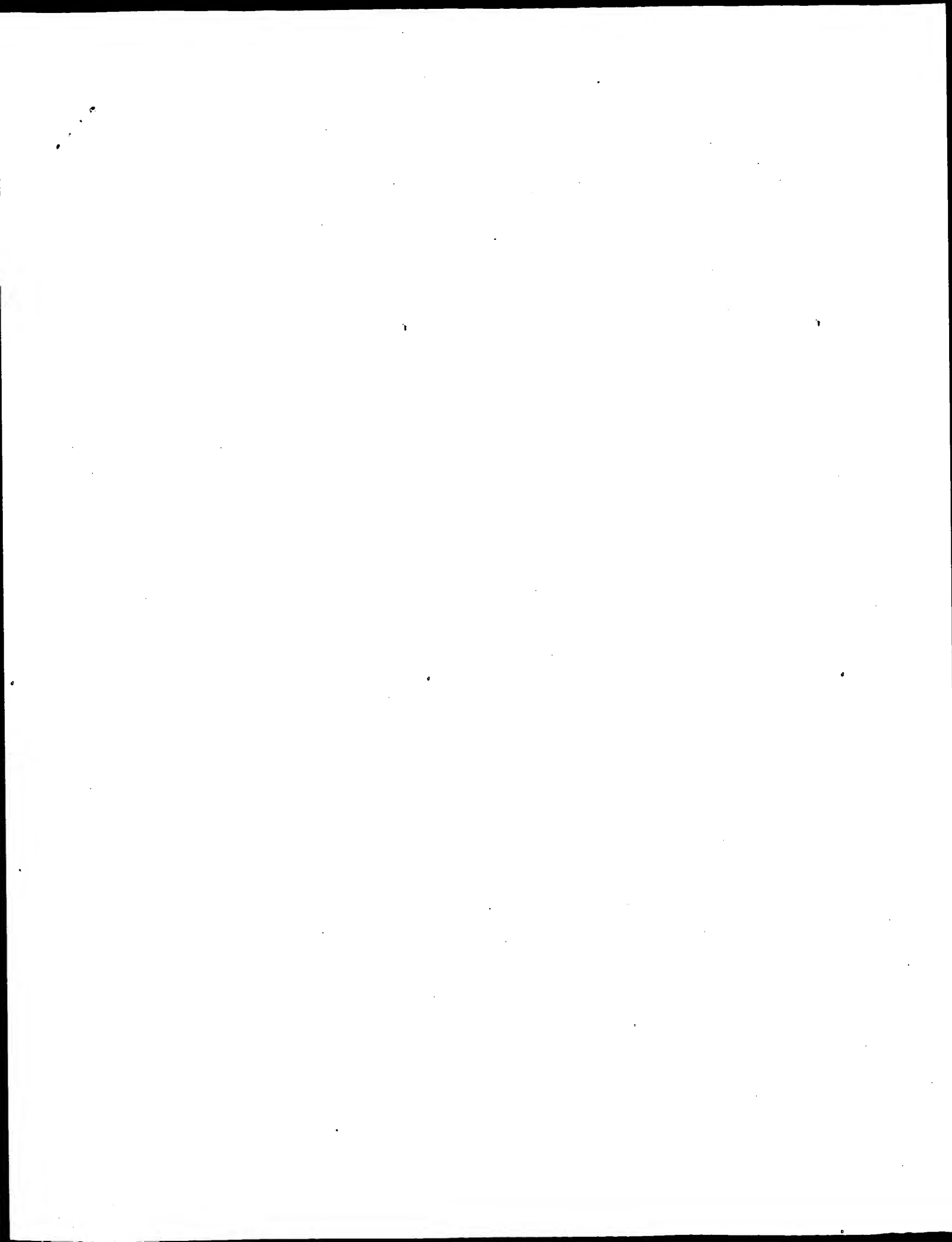
CC The present invention describes a method for chemically synthesizing a
CC peptide by solid phase synthesis. The method comprises processing a
CC carrier resin to which the C-terminal amino acid derivative of the
CC target peptide has been preliminarily bonded is characterized by adding
CC a protective group to inactivate a part of alpha-amino groups of an
CC amino acid capable of undergoing a peptide extension reaction on the
CC carrier resin. The method is useful for synthesizing a peptide by solid
CC phase synthesis. Long chain peptides can be readily synthesised in with
CC a high success ratio without using the segment condensation method or
CC gene manipulations. The present sequence represents an amino acid
CC peptide sequence given in an example from the present invention.

SQ Sequence 95 AA;

Query Match 21.6%; Score 120.5; DB 21; Length 95;
Best Local Similarity 52.0%; Pred. No. 7.2e-08;
Matches 26; Conservative 3; Mismatches 2; Indels 19; Gaps 1;

QY 1 MKVTVAENQFGPNRRVFIKRVSNVTHGRRIDIFASKNFHLQKNTLTGTR 50
DB 27 MKVTVAENQFGP-----DIFASKNFHLQKNTLTGSK 57

Search completed: April 20, 2003, 13:06:03
Job time : 128.382 secs



GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 18:04:50 ; Search time 17 Seconds

(without alignments)
757.765 Million cell updates/sec

Title: US-09-142-524D-3

Sequence: 1 MKVTVAFNQFGNRRVFIKR.....KFIIRVDGIAMYNPASWK 134

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	188	26.4	374	2 JC2124	major allergen Cry
2	188	26.4	374	2 JC2123	major allergen Cry
3	148.5	20.9	514	2 S48730	Cry j II protein -
4	148.5	20.9	514	2 JC2498	second major aller
5	118.5	16.7	514	2 JC7100	polygalacturonase
6	110.5	15.5	542	2 TC6728	pectate lyase (EC
7	107.5	15.1	507	2 TC3666	Jun a 2 protein -
8	104.5	14.7	463	2 T46165	pectate lyase-like
9	103	14.5	404	2 S12209	pectate lyase (EC
10	99	13.9	472	2 T51456	pectate lyase-like
11	98.5	13.9	396	2 A39099	allergen Amb a 1.1
12	97	13.6	397	2 C53240	allergen Amb a 1.3
13	97	13.6	397	2 C39099	allergen Amb a 1.3
14	97	13.6	455	2 T00856	pectate lyase (EC
15	95.5	13.4	459	2 T05556	pectate lyase (EC
16	94.5	13.3	404	2 T05556	hypothetical prote
17	94.5	13.3	431	2 F86179	hypothetical prote
18	93	13.1	397	2 S26211	pectate lyase (EC
19	93	13.1	398	2 B39099	allergen Amb a 1.2
20	93	13.1	450	2 T09524	allergen Amb a 1.2
21	92	12.9	397	2 E53240	allergen Amb a 1.1
22	92	12.9	398	2 B53240	allergen Amb a 1.1
23	91.5	12.9	368	2 G86427	probable pectate 1
24	91	12.8	398	2 T07058	pectate lyase (EC
25	89	12.5	392	2 D53240	allergen Amb a 1.4
26	89	12.5	434	2 S29612	pectate lyase (EC
27	88	12.4	438	2 S4335	pectate lyase (EC
28	84	11.8	374	2 T05240	pectate lyase (EC
29	84	11.8	374	2 H85148	probable pectate 1

30	84	11.8	390	2 H86253	hypothetical prote
31	84	11.6	394	2 T49115	pectate lyase like
32	82.5	11.6	449	2 S27098	pectate lyase (EC
33	82	11.5	394	2 T49116	pectate lyase like
34	81	11.4	219	2 C84647	probable synaptobr
35	79	11.1	274	2 T03932	pectate lyase (EC
36	78.5	11.0	418	2 T07701	pectate lyase (EC
37	78	11.0	341	2 T47653	pectate lyase-like
38	76.5	10.8	648	2 S10869	enterotoxin A - Cl
39	76	10.7	398	2 F72335	hypothetical prote
40	74.5	10.5	368	2 F83785	required for glyco
41	74	10.4	219	2 T04630	synaptobrevin homo
42	71.5	10.1	657	2 T22451	hypothetical prote
43	71	10.0	2710	2 A37652	coxin A - Clostrid
44	70.5	9.9	622	2 S17402	paraspinal crystal
45	70.5	9.9	1003	2 T26746	hypothetical prote

ALIGNMENTS

RESULT 1
JC2124
major allergen Cry j I precursor (clone pCCI-15) - Japanese cedar
C:Species: Cryptomeria japonica (Japanese cedar)
C>Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 21-Jul-2000
C/Accession: JC2124
R/Sone, T.; Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.
Biochem. Biophys. Res. Commun. 199, 619-625, 1994
A/Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japanese
A/Reference number: JC2123; NCID:94183234; PMID:8135802
A/Accession: JC2124
A/Molecule type: mRNA
A/Residues: 1-374 <SON>
A/Cross-references: GB:D26545; NID:G493633; PIDN:BA0543.1; PID:G493634
A/Experimental source: pollen
A/Note: the authors described carbohydrate binding site for residue 279
C/Keywords: pectate lyase, LAR59
C/Keywords: glycoprotein; pollen
F/1-21/Domain: signal sequence #status predicted <SIG>
F/22-374/Product: major allergen Cry j I (clone pCCI-15) #status predicted <MAT>
F/158,191,293,354/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.4%; Score 188; DB 2; Length 374;
Best Local Similarity 71.2%; Pred. No. 1.9e-12;
Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 73 RPLWIFSGNMNITKLMPTIAGYKTFDGRRAEVSVHVNGAKFIRVDGII 124
Db 87 RPLWIFSGNMNITKLMPTIAGYKTFDGRRAEVSVHVNGAKFIRVDGII 138

RESULT 2
JC2123
major allergen Cry j I precursor (clone pCCI-2-2) - Japanese cedar
C:Species: Cryptomeria japonica (Japanese cedar)
C>Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 21-Jul-2000
C/Accession: JC2123; PC2065
R/Sone, T.; Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.
Biochem. Biophys. Res. Commun. 199, 619-625, 1994
A/Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japanese
A/Reference number: JC2123; NCID:94183234; PMID:8135802
A/Accession: JC2123
A/Molecule type: mRNA
A/Residues: 1-374 <SON>
A/Cross-references: GB:D26544; NID:G493633; PIDN:BA0543.1; PID:G493632
A/Experimental source: pollen
A/Accession: PC2065
A/Molecule type: protein
A/Residues: 22-53;58-81;219-232;236-258;299-307;346-372 <SQ2>
A/Note: the authors described carbohydrate binding site for residue 279
C/Superfamily: pectate lyase LAR59
C/Keywords: glycoprotein; pollen

F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-374/Product: major allergen Cry j I (clone pCCI-2-2) #status predicted <MAT>
F:158,191,293,354/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.4%; Score 188; DB 2; Length 374;
Best Local Similarity 71.2%; Pred. No. 1.9e-12;
Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 73 RPLMIIFSGNNIKLKMPMTIAGYKTFDGRRAEVSIVHNGAKFIRVDGI 124
DB 87 RPLMIIFSGNNIKLKMPMTIAGYKTFDGRAGYVIGNGPCVFIRKSVNI 138

RESULT 3

S48730
Cry j II protein - Japanese cedar
C:Species: Cryptomeria japonica (Japanese cedar)
C:Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 28-May-1999
C:Accession: S48730
R:Namba, M.; Kurose, M.; Toriue, K.; Hino, K.; Taniguchi, Y.; Fukuda, S.; Usui, M.; Kur
FEBB Lett. 353, 124-128, 1994
A:Title: Molecular cloning of the second major allergen, Cry j II, from Japanese cedar
A:Reference number: S48730; MUID:95010777; PMID:7926035
A:Accession: S48730
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-514 <NM>
A:Cross-references: GB:D37765; NID:9577695; PIDN:BA07021.1; PID:dl007598; PID:9577696

Query Match 20.9%; Score 148.5; DB 2; Length 514;
Best Local Similarity 41.4%; Pred. No. 4.9e-08;
Matches 41; Conservative 8; Mismatches 21; Indels 29; Gaps 2;

QY 31 IDIFASKNFHLQKNTIGTGRKKNRIMLOPAKLTGFTLMGRPLMIIFSGNNIKLKMP 90
DB 236 IDIFASKNFHLQKNTIGT-----DDCVAIIGTSSNIYIEDL 272

QY 91 MYINGKTFD-----RRAEVSIVHNGAKFIRVDGI 123
DB 273 ICGPHGISIGSLGRNSRAEVSIVHNGAKFIDTONGL 311

RESULT 4

JC2498
second major allergen Cry j II precursor - Japanese cedar
C:Species: Cryptomeria japonica (Japanese cedar)
C:Date: 16-Mar-1995 #sequence_revision 26-May-1995 #text_change 21-Jul-2000
C:Accession: JC2498; PC2346; A60147

R:Komiyama, N.; Sone, T.; Shimizu, K.; Morikubo, K.; Kino, K.
Biochem. Biophys. Res. Commun. 201, 1021-1028, 1994
A:Title: cDNA cloning and expression of Cry j II, the second major allergen of Japanese
A:Reference number: JC2498; MUID:94271186; PMID:8002972
A:Accession: JC2498
A:Molecule type: mRNA
A:Residues: 1-514 <XM>
A:Cross-references: DDBJ:D29772; NID:9506857; PIDN:BA06172.1; PID:9506858

A:Accession: PC2346
A:Molecule type: protein
A:Residues: 52-61 <KO2>
R:Sakaguchi, M.; Inouye, S.; Tanai, M.; Ando, S.; Usui, M.; Matuhasi, T.
Allergy 45, 309-312, 1990
A:Title: Identification of the second major allergen of Japanese cedar pollen.
A:Reference number: A60147; MUID:90342988; PMID:2382797
A:Accession: A60147
A:Molecule type: protein
A:Residues: 55-64 <SAK>
C:Keywords: glycoprotein; pollen

F:1-54/Domain: signal sequence #status predicted <SIG>
F:55-460/Product: second major allergen Cry j #status predicted <MAT>
F:429,460,472/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 20.9%; Score 148.5; DB 2; Length 514;
Best Local Similarity 41.4%; Pred. No. 4.9e-08;

Matches 41; Conservative 8; Mismatches 21; Indels 29; Gaps 2;

QY 31 IDIFASKNFHLQKNTIGTGRKKNRIMLOPAKLTGFTLMGRPLMIIFSGNNIKLKMP 90
DB 236 IDIFASKNFHLQKNTIGT-----DDCVAIIGTSSNIYIEDL 272

QY 91 MYINGKTFD-----RRAEVSIVHNGAKFIRVDGI 123
DB 273 ICGPHGISIGSLGRNSRAEVSIVHNGAKFIDTONGL 311

RESULT 5

JC7100
polygalacturonase Cha o 2 - Japanese cypress
C:Species: Chamaecyparis obtusa (Japanese cypress)
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
C:Accession: JC7100; PC7026
R:Mori, T.; Yokoyama, M.; Komiyama, N.; Okano, M.; Kino, K.
Biochem. Biophys. Res. Commun. 263, 166-171, 1999

A:Title: Purification, identification, and cDNA cloning of Cha o 2, the second major all
A:Reference number: JC7100; MUID:9417540; PMID:10486272
A:Accession: JC7100
A:Molecule type: protein
A:Residues: 1-514 <MO>
A:Accession: PC7026
A:Molecule type: protein
A:Residues: 51-62 <MO2>

Query Match 16.7%; Score 118.5; DB 2; Length 514;
Best Local Similarity 35.4%; Pred. No. 8.3e-05;
Matches 35; Conservative 10; Mismatches 25; Indels 29; Gaps 2;

QY 31 IDIFASKNFHLQKNTIGTGRKKNRIMLOPAKLTGFTLMGRPLMIIFSGNNIKLKMP 90
DB 236 IDIFASKNFHLQKNTIGT-----DDCVAIIGTSSNIYIEDL 272

QY 91 MYINGKTFD-----DGRRAEVSIVHNGAKFIRVDGI 123
DB 273 ICGPHGISIGSLGRNSRAEVSIVHNGAKFIDTONGL 311

RESULT 6

T06728
pectate lyase (EC 4.2.2.2) F28P10.100 - Arabidopsis thaliana
N:Alternate names: protein F28P10.100
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C:Accession: T06728

R:Quellier, F.; Choline, N.; Robert, C.; Broctier, P.; Wincker, P.; Cactolico, L.; Artigu
submitted to the Protein Sequence Database, April 1999
A:Reference number: Z15793
A:Accession: T06728
A:Molecule type: DNA
A:Residues: 1-542 <OUE>
A:Cross-references: EMBL:AL049655

A:Experimental source: cultivar Columbia; BAC clone F28P10
C:Genetics:
A:Map position: 3
A:Introns: 46/2; 346/3; 413/2; 480/2
A:Note: F28P10.100
C:Superfamily: pectate lyase LAR59
C:Keywords: carbon-oxygen lyase

Query Match 15.5%; Score 110.5; DB 2; Length 542;
Best Local Similarity 43.3%; Pred. No. 0.00064;
Matches 26; Conservative 9; Mismatches 20; Indels 5; Gaps 2;

QY 67 FTLMGRRPLMIIFSGNNIKLKMPMTIAGYKTFDGRRAEVSIVHNGAKFIRVDGI 124
DB 139 YGVIOEPLMTIVFSSNMLIRLKHILINSYTLDRG---SAVHTNGCLTLQYVHHI 195

RESULT 7

UC7366
 Jun a 2 protein - mountain cedar
 C/Species: Juniperus ashei (mountain cedar)
 C/Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000
 C/Accession: UC7366, PC7093
 R/Accession: M.; Miyahara, M.; Shimizu, K.; Kino, K.; Tsumoto, H.
 Biochem. Biophys. Res. Commun. 275, 195-202, 2000
 A/Title: Purification, identification, and cDNA cloning of Jun a 2, the second major all
 A/Reference number: UC7366
 A/Accession: UC7366
 A/Molecule type: mRNA
 A/Residues: 1-507 <YOK>
 A/Cross-references: GB:AJ404653
 A/Accession: PC7093
 A/Molecule type: protein
 A/Residues: 55-63 <Y02>
 C/Comment: This protein, a second major allergen of mountain cedar pollen, which is inv
 C/Keywords: glycoprotein; pollen

Query Match 15.1%; Score 107.5; DB 2; Length 507;
 Best Local Similarity 32.3%; Pred. No. 0.0012;
 Matches 32; Conservative 13; Mismatches 25; Indels 29; Gaps 2;
 Oy 31 IDIASKNFHLQKNTIGTGRKKNRIMLOPAKLTGFTLMGRRLWIFSGNNNTKMKMP 90
 Db 237 IDIASKNFHLQKNTIGTGRKKNRIMLOPAKLTGFTLMGRRLWIFSGNNNTKMKMP 90
 Oy 91 MYIAGYKTFDQ-----RRAEVSYVHNGAKETRRVDGI 123
 Db 274 TCGPHGMSISGLCKGSRSEVSRVHLDGAKFIDTQNGL 312

RESULT 8
 T46165
 pectate lyase-like protein - Arabidopsis thaliana
 N/Alternate names: protein T46120
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 02-Sep-2000
 C/Accession: T46165
 R/Yakutara, G.; Fartmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Weichselgartner, M.;
 submitted to the Protein Sequence Database, December 1999
 A/Reference number: Z23025
 A/Accession: T46165
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-463 <NTA>
 A/Cross-references: EMBL:AL132958
 A/Experimental source: cultivar Columbia; BAC clone T4D2
 C/Genetics:
 A/Map position: 3
 A/Intons: 27/2; 112/2; 182/1; 265/3; 329/3; 399/2
 A/Note: T4D2.120
 C/Superfamily: pectate lyase LAT59

Query Match 14.7%; Score 104.5; DB 2; Length 463;
 Best Local Similarity 30.5%; Pred. No. 0.0024;
 Matches 32; Conservative 8; Mismatches 32; Indels 33; Gaps 3;
 Oy 52 WKNRRLW-----QFAKLTLG-----FTLMGRRLWIF 79
 Db 66 WPNRQGLADCGIGFGVALGCKGQFYVTDSDDAVNPGRITRGVIOEPLWIF 125
 Oy 80 SGNNNTKMKMPYIAGYKTFDGRRAEVSYVHNGAKETRRVDGI 124
 Db 126 PSNNNTKMKMPYIAGYKTFDGRRAEVSYVHNGAKETRRVDGI 169

RESULT 9
 S12209
 pectate lyase (EC 4.2.2.2) - tomato
 C/Species: Lycopersicon esculentum (tomato)
 C/Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 23-Jul-1999

C/Accession: S12209
 R/Budelier, K.A.; Smith, A.G.; Gasser, C.S.
 Mol. Gen. Genet. 224, 183-192, 1990
 A/Title: Regulation of a stylar transmitting tissue-specific gene in wild-type and trans
 A/Reference number: S12209; MUID:91117185; PMID:2277637
 A/Accession: S12209
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-404 <BDU>
 A/Cross-references: GB:X55193; NID:919161; PID:CA38979.1; PID:919162
 C/Superfamily: pectate lyase LAT59
 C/Keywords: carbon-oxygen lyase

Query Match 14.5%; Score 103; DB 2; Length 404;
 Best Local Similarity 30.8%; Pred. No. 0.0029;
 Matches 33; Conservative 13; Mismatches 31; Indels 30; Gaps 4;
 Oy 28 GRRIDPASKNFHLQKNTIGTGRKKNRIM-----LOPAKLTGFTLM 70
 Db 67 GKSPAFSTICAIQFGKNAIG---GKNRIVVTDSGNDPVPKPTLRHA-----VI 116
 Oy 71 GRPPLWIFSGNNNTKMKMPYIAGYKTFDGRRAEVSYVHNGAKFI 117
 Db 117 QDEPLWIFGRDMVITQKQELWMSYKTIIDRGAS---VHISGPCI 160

RESULT 10
 T51456
 pectate lyase-like protein - Arabidopsis thaliana
 N/Alternate names: protein F2G14.230
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 17-Nov-2000
 C/Accession: T51456
 R/Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
 submitted to the Protein Sequence Database, August 2000
 A/Reference number: Z25394
 A/Accession: T51456
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-472 <SAT>
 A/Cross-references: EMBL:AL391146
 A/Experimental source: cultivar Columbia; BAC clone F2G14
 C/Genetics:
 A/Map position: 5
 A/Intons: 65/2; 316/3; 398/3
 A/Note: F2G14_230
 C/Superfamily: pectate lyase LAT59

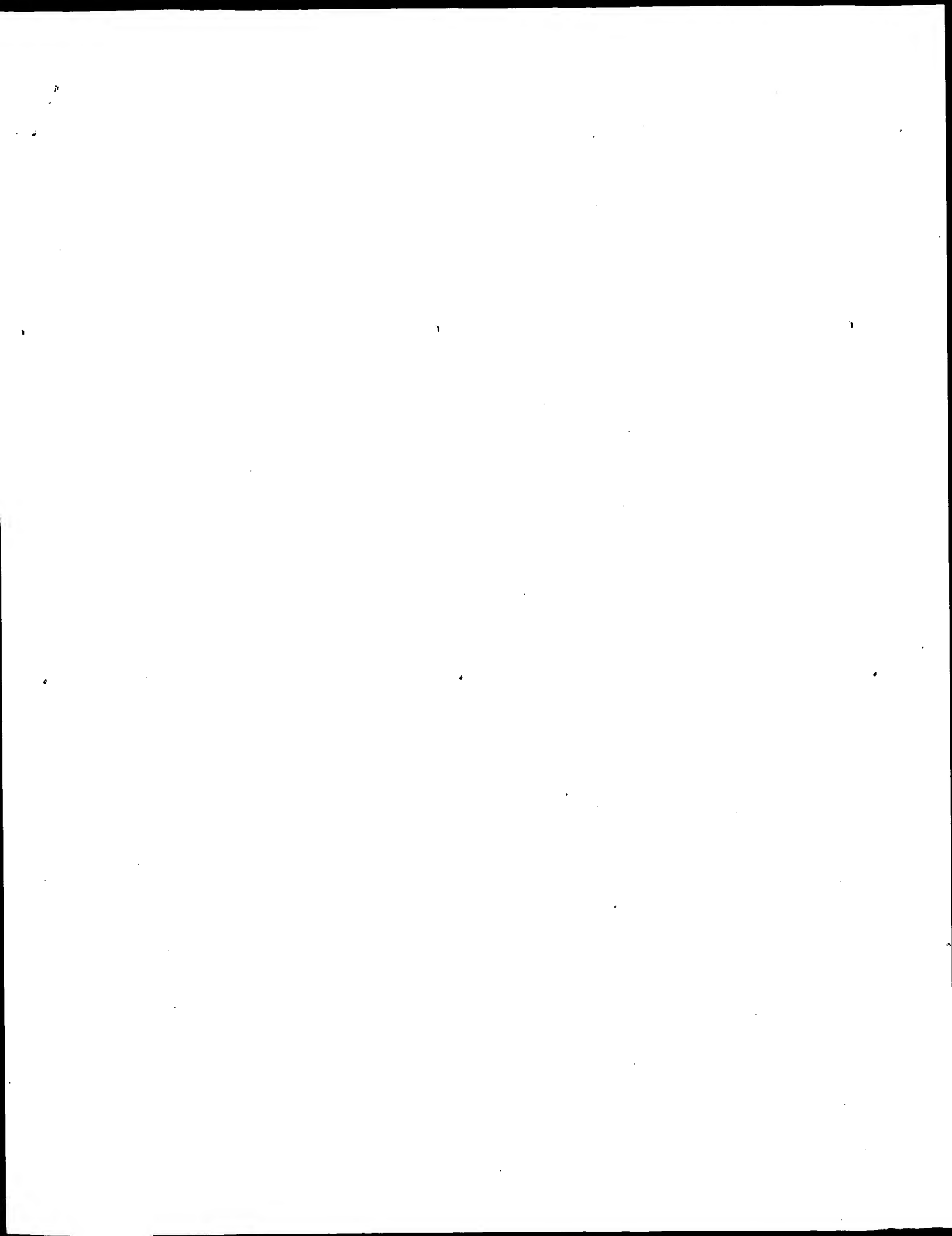
Query Match 13.9%; Score 99; DB 2; Length 472;
 Best Local Similarity 28.2%; Pred. No. 0.0094;
 Matches 29; Conservative 17; Mismatches 41; Indels 16; Gaps 3;
 Oy 34 FASKNFHLQKNTIGTGR---WKNRRLWLOPAKLT-----TGFTLMGRRLWIF 79
 Db 134 WAKRRKLTLCVSGFGHRTTGKRGRIYVTSNLDMDVNPGRITRGVIOEPLWIF 193
 Oy 80 SGNNNTKMKMPYIAGYKTFDGRRAEVSYVHNG--AKTIRRV 120
 Db 194 KNDWSIRLNQELLINSHKTIIDRGANVHVAHGATMGVKNV 236

RESULT 11
 A39099
 allergen Amb a 1.1 precursor - common ragweed
 C/Species: Ambrosia artemisiifolia (common ragweed)
 C/Date: 27-Nov-1991 #sequence_revision 03-Apr-1992 #text_change 29-Sep-1999
 C/Accession: A39099; A60895; A53240
 R/Rafnar, T.; Griffith, I.J.; Kuo, M.; Bond, J.F.; Rogers, B.L.; Klapper, D.G.
 J. Biol. Chem. 266, 1229-1236, 1991
 A/Title: Cloning of Amb a 1 (antigen E), the major allergen family of short ragweed pol
 A/Reference number: A39099; MUID:91093235; PMID:1702434
 A/Accession: A39099
 A/Molecule type: mRNA

A/Residues: 1-396 <RAF>
 A/Cross-references: GB:M63116
 R/Smith, J.J.; Olson, J.R.; Klapper, D.G.
 Mol. Immunol. 25, 355-365, 1988
 A/Title: Monoclonal antibodies to denatured ragweed pollen allergen Amb a I: characterized a I.
 A/Reference number: A60895; MUID:88288254; PMID:2456454
 A/Accession: A60895
 A/Molecule type: protein
 A/Residues: 256-273/292-303, 'W', 305-306 <SMT>
 R/Griffith, I.J.; Pollock, J.; Klapper, D.G.; Rogers, B.L.; Nault, A.K.
 Int. Arch. Allergy Appl. Immunol. 96, 296-304, 1991
 A/Title: Sequence polymorphism of Amb a I and Amb a II, the major allergens in Ambrosia
 A/Reference number: A53240; MUID:92234570; PMID:1809687
 A/Accession: A53240
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-396 <GRI>
 A/Cross-references: GB:M60558; NID:G166434; PIDN:AAA32665.1; PID:G166435
 C/Superfamily: pectate lyase LAT59
 C/Keywords: glycoprotein, pollen
 F/36/Binding site: carbohydrate (Asn) (covalent) #status predicted
 Query Match 13.6%; Score 97; DB 2; Length 396;
 Best Local Similarity 34.7%; Pred. No. 0.0087;
 Matches 26; Conservative 12; Mismatches 22; Indels 15; Gaps 3;
 RESULT 12
 CS3240
 allergen Amb a I.3 precursor - common ragweed
 C/Species: Ambrosia artemisiifolia (common ragweed)
 C/Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 18-Sep-1998
 C/Accession: CS3240
 R/Griffith, I.J.; Pollock, J.; Klapper, D.G.; Rogers, B.L.; Nault, A.K.
 Int. Arch. Allergy Appl. Immunol. 96, 296-304, 1991
 A/Title: Sequence polymorphism of Amb a I and Amb a II, the major allergens in Ambrosia
 A/Reference number: A53240; MUID:92234570; PMID:1809687
 A/Accession: CS3240
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-397 <GRI>
 A/Cross-references: GB:M80560
 C/Superfamily: pectate lyase LAT59
 C/Keywords: pollen
 Query Match 13.6%; Score 97; DB 2; Length 397;
 Best Local Similarity 29.5%; Pred. No. 0.013;
 Matches 31; Conservative 7; Mismatches 33; Indels 34; Gaps 3;
 RESULT 13
 CS3909
 allergen Amb a I.3 - common ragweed
 C/Species: Ambrosia artemisiifolia (common ragweed)
 C/Date: 27-Nov-1991 #sequence_revision 03-Apr-1992 #text_change 29-Sep-1999
 C/Accession: C39099

R/Rafnar, T.; Griffith, I.J.; Kuo, M.; Bond, J.F.; Rogers, B.L.; Klapper, D.G.
 J. Biol. Chem. 266, 1229-1236, 1991
 A/Title: Cloning of Amb a I (antigen E), the major allergen family of short ragweed poll
 A/Reference number: A39099; MUID:91092325; PMID:1702434
 A/Accession: C39099
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-397 <RAF>
 A/Cross-references: GB:M62961; NID:G166440; PIDN:AAA32668.1; PID:G166441
 C/Superfamily: pectate lyase LAT59
 C/Keywords: pollen
 Query Match 13.6%; Score 97; DB 2; Length 397;
 Best Local Similarity 29.5%; Pred. No. 0.013;
 Matches 31; Conservative 7; Mismatches 33; Indels 34; Gaps 3;
 RESULT 14
 T00856
 pectate lyase (EC 4.2.2.2) T20F6.14 - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 02-Feb-2001
 C/Accession: T00856; A84440
 R/Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
 Submitted to the EMBL Data Library, March 1998
 A/Description: Arabidopsis thaliana chromosome II BAC T20F6 genomic sequence.
 A/Reference number: Z14206
 A/Accession: T00856
 A/Status: translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-455 <ROU>
 A/Cross-references: EMBL:AC002521; NID:G2947056; PIDN:AA05350.1; PID:G2947069
 A/Experimental source: cultivar Columbia
 R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A/Reference number: A84420; MUID:20083487; PMID:10617197
 A/Accession: A84440
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-455 <STO>
 A/Cross-references: GB:AE002093; NID:G2947069; PIDN:AA05350.1; GSPDB:GN00139
 C/Genetics:
 A/Map position: 2
 A/Insertions: 66/2; 295/3; 376/3
 A/Note: T20F6.14
 C/Superfamily: pectate lyase LAT59
 C/Keywords: carbon-oxygen lyase
 Query Match 13.6%; Score 97; DB 2; Length 455;
 Best Local Similarity 46.2%; Pred. No. 0.015;
 Matches 24; Conservative 9; Mismatches 13; Indels 6; Gaps 3;
 RESULT 15
 G86278
 hypothetical protein F14I7.19 (imported) - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 01-Mar-2002
C:Accession: G862278
R:Theologidis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Corn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.,
Nansen, N.P.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.D.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
Rizzo, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzall,
A.; Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yi, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: G86278
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-459 <STO>
A:Cross-references: GB:AE005172; NID:g7262684; PIDN:AA43942.1; GSPDB:GN00141
C:Genetics:
C:Map position:
C:Superfamily: pectate lyase LAM59



GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 18:04:50 ; Search time 13 Seconds

(without alignments)
427.525 Million cell updates/sec

Title: US-09-142-524D-3

Perfect score: 711
Sequence: 1 MKVTAFNQGPNRRVFIK.....KPIRRVGIILAAQNDPSWK 134

Scoring table:

BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	188	26.4	374	1	SBP_CRYJA
2	150	21.1	367	1	MPA1_JUNAS
3	148.5	20.9	514	1	MPA2_CRYJA
4	143	20.1	346	1	MPA1_CUPAR
5	142	20.0	375	1	MPA1_CHAOS
6	103	14.5	404	1	MPA1_LYCES
7	98.5	13.9	396	1	MPA1_AMBAR
8	97	13.6	397	1	MPA1_AMBAR
9	93	13.1	397	1	PEL_TOBAC
10	93	13.1	398	1	MPA2_AMBAR
11	92	12.9	397	1	MPA2_AMBAR
12	91	12.8	398	1	PESE_LYCES
13	89	12.5	392	1	MPA1_AMBAR
14	89	12.5	434	1	PEL_LILLO
15	82.5	11.6	449	1	PESE_LYCES
16	80.5	11.5	234	1	A298_DROME
17	80.5	11.3	605	1	PHAC_METEX
18	71	10.0	2710	1	TOXA_CLODI
19	70.5	9.9	622	1	C2AC_BACTU
20	69.5	9.8	338	1	YUX8_YEAST
21	68.5	9.6	234	1	A298_DROSI
22	68.5	9.6	551	1	CHIT_NPVAC
23	67.5	9.5	985	1	AGLU_ASFOR
24	67.5	9.5	1070	1	YHVA_YEAST
25	67	9.4	757	1	YHVA_YEAST
26	66.5	9.4	110	1	Y12K_SMSV4
27	66	9.3	375	1	Y4ED_RHSN
28	66	9.3	382	1	PHYT_BACST
29	66	9.3	646	1	NODQ_RHSB
30	65	9.1	576	1	DPEP_SOLIS
31	65	9.1	726	1	YBID_SCHPO
32	65	9.1	781	1	TL2D_CHICK
33	64.5	9.1	1010	1	CLPP_CHLEU

34	64	9.0	532	1	PUR2_ARATH	P52420 arabidopsis
35	64	9.0	633	1	NODQ_RHSB	O07309 r nodd bifu
36	63.5	8.9	196	1	YMO7_YEAST	Q04487 saccharomyc
37	63.5	8.9	309	1	Y4OQ_RHSN	P55602 rhizobium s
38	63.5	8.9	326	1	SCER_PEDPE	P43472 pedicoccus
39	63	8.9	347	1	NI2M_RABIT	O79428 oryctolagus
40	63	8.9	477	1	GIGA_SYNT3	P74521 synechocyst
41	63	8.9	522	1	CPV1_ORENT	P70091 oreochromis
42	63	8.9	725	1	GUNG_CLOCE	P37700 clostridium
43	63	8.9	755	1	COMP_RAT	P35444 rattus norv
44	63	8.9	1121	1	WDRE_HUMAN	O9nnw5 homo sapien
45	62.5	8.8	262	1	FLGG_AGRIS	Q44338 agrobacteri

ALIGNMENTS

RESULT 1
ID SBP_CRYJA STANDARD, PRT, 374 AA.
AC P18632;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Sugt basic protein precursor (SBP) (Major allergen Cry j 1) (Cry j 1).
OS Cryptomeria japonica (Japanese cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.
OK NCBI_TaxID=3369;
[1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Pollen;
RX MEDLINE=94183234; PubMed=6135802;
RA Some T., Komiyama N., Shimizu K., Kusakabe T., Morikubo K.,
RA Kino K.;
RT "Cloning and sequencing of cDNA coding for Cry j I, a major allergen
RT of Japanese cedar pollen.";
RL Biochem. Biophys. Res. Commun. 199:619-625(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pollen;
RX Namba M., Kurose M., Torigoe K., Fukuda S., Kurimoto M.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 22-41.
RC TISSUE=Pollen;
RX MEDLINE=89031257; PubMed=3181436;
RA Tanai M., Ando S., Usui M., Kurimoto M., Sakaguchi M., Inoue S.,
RA Matsumoto T.;
RT "N-terminal amino acid sequence of a major allergen of Japanese cedar
RT pollen (Cry j I).";
RL FEBS Lett. 239:329-332(1988).
RN [4]
RP CARBOHYDRATES.
RC TISSUE=Pollen;
RX MEDLINE=95003748; PubMed=7920021;
RA Hijikata A., Matsumoto I., Kojima K., Ogawa H.,
RT "Antigenicity of the oligosaccharide moiety of the Japanese cedar
RT (Cryptomeria japonica) pollen allergen, Cry j I.";
RL Int. Arch. Allergy Immunol. 105:198-202(1994).
RN [5]
RP STRUCTURE OF CARBOHYDRATES.
RC TISSUE=Pollen;
RX MEDLINE=95332249; PubMed=7608114;
RA Hino K., Yamamoto S., Sano O., Taniguchi Y., Kohno K., Usui M.,
RT Fukuoka S., Hanawa H., Haruyama H., Kurimoto M.;
RT "Carbohydrate structures of the glycoprotein allergen Cry j I from
RT Japanese cedar (Cryptomeria japonica) pollen.";
RL J. Biochem. 117:289-295(1995).
CC -1- PPM: CONTAINS FUOSE/YULOSE-CONTAINING N-LINKED OLIGOSACCHARIDES.
CC -1- DISEASE: THIS PROTEIN IS THE MAJOR ALLERGEN OF JAPANESE CEDAR
CC POLLEN, THE MOST COMMON POLLEN ALLERGEN IN JAPAN.
CC -1- MISCELLANEOUS: THE SEQUENCE OF CRY J I FORM A IS SHOWN HERE. FORM

B DIFFERS IN SIX POSITIONS.
 -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 AMB A I/AMB A II/CRY J I SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: D26544; BAA05542.1; -
 DR EMBL: D26545; BAA05543.1; -
 DR EMBL: D34639; BAA07020.1; -
 DR PIR: A44773; A44773.
 DR GlycoSuiteDB; P18632; -
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec lyase; 1.
 DR PRINTS; PR00807; AMBALLALRGEN.
 KW Allergen; Glycoprotein; Multigene family; Signal.
 FT SIGNAL 1 21
 FT CHAIN 1 21
 FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (COMPLEX).
 FT CARBOHYD 293 293 /FTIG-CAR 000135.
 FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (COMPLEX).
 FT CARBOHYD 354 354 /FTIG-CAR 000136.
 FT VARIANT 12 12 L -> P (IN CRY J 1-B).
 FT VARIANT 143 143 H -> Y (IN CRY J 1-B).
 FT VARIANT 202 202 S -> T (IN CRY J 1-B).
 FT VARIANT 221 221 L -> S (IN CRY J 1-B).
 FT VARIANT 358 358 K -> H (IN CRY J 1-B).
 FT VARIANT 361 361 Q -> Q (IN CRY J 1-B).
 FT VARIANT 361 361 Q -> Q (IN CRY J 1-B).
 SQ SEQUENCE 374 AA; 40645 MW; 74AB25950248F56F CRC64;

Query Match 26.4%; Score 188; DB 1; Length 374;
 Best Local Similarity 71.2%; Pred. No. 2.7e-12;
 Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

 OY 73 RPLMIIFSGMMNLIKMPMTIAGKTFDGRRAVSYYVHNGAKFIRVDGII 124
 DB 87 RPLMIIFSGMMNLIKMPMTIAGKTFDGRRAVSYYVHNGAKFIRVDGII 138

RESULT 2
 MPAL JUNAS STANDARD; PRT; 367 AA.
 AC P81254; O9ZNU7;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Major pollen allergen Jun a 1 precursor.
 OS Juniperus ashei (Osark white cedar).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
 OC NCBI_TaxID=13101;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 31-38; 42-50; 58-80; 88-94;
 RP 117-124; 134-140; 160-164; 256-263 AND 322-325.
 RC TISSUE=Pollen;
 RX MEDLINE=99414163; PubMed=10482836;
 RA Madero-Horintu T.M., Goldblum R.M., Kurosky A., Wood T.G.,
 RA Brooks E.G.;
 RT "Molecular cloning of mountain cedar (Juniperus ashei) pollen major
 RT allergen, Jun a 1.";
 RT J. Allergy Clin. Immunol. 104:613-617(1999).
 RN [2]
 RP SEQUENCE OF 22-50.
 RP TISSUE=Pollen;
 RX MEDLINE=99414162; PubMed=10482835;
 RA Madero-Horintu T., Goldblum R.M., Kurosky A., Goetz D.W.,

RA Brooks E.G.;
 RT "Isolation and characterization of the mountain cedar (Juniperus
 RT ashei) pollen major allergen, Jun a 1.";
 RT J. Allergy Clin. Immunol. 104:608-612(1999).
 RL [1]
 CC -1- DISORDER: THIS PROTEIN CAUSES SEVERAL SEASONAL ALLERGIC RHINITIS
 CC IN NORTH AMERICA.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: AFI06663; AAD03609.1; -
 DR EMBL: AFI06662; AAD03608.1; -
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec lyase; 1.
 DR PRINTS; PR00807; AMBALLALRGEN.
 KW Allergen; Glycoprotein; Signal.
 FT SIGNAL 1 21
 FT CHAIN 1 21
 FT CARBOHYD 148 148 MAJOR POLLEN ALLERGEN JUN A 1.
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 367 AA; 39824 MW; FC9B81E675652E49 CRC64;

Query Match 21.1%; Score 150; DB 1; Length 367;
 Best Local Similarity 55.8%; Pred. No. 2.4e-08;
 Matches 29; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

 OY 73 RPLMIIFSGMMNLIKMPMTIAGKTFDGRRAVSYYVHNGAKFIRVDGII 124
 DB 87 KALMIIFSGMMNLIKMPMTIAGKTFDGRRAVSYYVHNGAKFIRVDGII 138

RESULT 3
 MPAL2 CRYVA STANDARD; PRT; 514 AA.
 AC P43212;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Possible polylacturonase precursor (EC 3.2.1.15) (PG) (Pectinase)
 DE (Major pollen allergen Cry j 2) (Cry j II).
 OS Cryptomeria japonica (Japanese cedar).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.
 OC NCBI_TaxID=3369;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RP TISSUE=Pollen;
 RX MEDLINE=95010777; PubMed=7926035;
 RA Namba M., Kurose M., Toriige K., Hino K., Taniguchi Y., Fukuda S.,
 RA Usui M., Kurimoto M.;
 RT "Molecular cloning of the second major allergen, Cry j II, from
 RT Japanese cedar pollen.";
 RT FEBS Lett. 353:124-128(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP TISSUE=Pollen;
 RX MEDLINE=94271186; PubMed=8002972;
 RA Komiyama N., Sone T., Shimizu K., Morikubo K., Kino K.;
 RT "cDNA cloning and expression of Cry j II the second major allergen of
 RT Japanese cedar pollen.";
 RT Biochem. Biophys. Res. Commun. 201:1021-1028(1994).
 RN [3]
 RP SEQUENCE OF 55-64.
 RP MEDLINE=90342988; PubMed=2382797;
 RA Sakaguchi M., Inouye S., Tanai M., Ando S., Usui M., Matuhasi T.;
 RT "Identification of the second major allergen of Japanese cedar

RT pollen." ;
 RL Allergy 45:309-312(1990).
 CC -1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-
 CC galactosiduronic linkages in pectate and other galacturonans.
 CC -1- SUBCELLULAR LOCATION: SECRETED OR AMYLOPLAST (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
 CC (POLYGALACTURONASES).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D37765; BAA07021.1; -;
 DR EMBL; D29772; BAA06172.1; -;
 DR HSRP; P26509; 1BHE.
 DR InterPro; IPR000743; GH28.
 DR Pfam; PF00295; Glyco_hydro_28; 1.
 DR PROSITE; PS00502; POLYGALACTURONASE; 1.
 KM Hydrolyase; Glycosidase; Cell wall; Signal; Zymogen; Fruit ripening;
 KM Amyloplast; Glycoprotein; Allergen.
 FT SIGNAL 1 22
 FT PROPEP 23 45
 FT CHAIN 46 433
 FT PROPEP 434 514
 FT ACT SITE 278 278
 FT CARBOHYD 460 460
 FT CARBOHYD 472 472
 FT CONFLICT 5 5
 FT CONFLICT 12 12
 FT CONFLICT 34 35
 FT CONFLICT 37 37
 FT CONFLICT 88 88
 FT CONFLICT 98 98
 FT CONFLICT 451 451
 FT CONFLICT 454 454
 FT CONFLICT 504 504
 FT CONFLICT 507 507
 SQ SEQUENCE 514 AA; 56645 MW; 62461103FAD6302 CRC64;
 Query Match 20.9%; Score 148.5; DB 1; Length 514;
 Best Local Similarity 41.4%; Pred. No. 4.9e-08;
 Matches 41; Conservative 8; Mismatches 21; Indels 29; Gaps 2;
 QY 31 IDIFASGNPHLOKNTTGTGRMKNNTWLOFALGTGTLGRKRPMTIFSGNNITLKM 90
 DB 236 IDIFASGNPHLOKNTTGTG-----DDCVAIGTGSNIVIEDL 272
 QY 91 MYIAGYKTFD-----RAEVSYYVHNGAKFIRRVGDI 123
 DB 273 ICGPHGISIGSLGRENSRAEVSYYVHNGAKFIDTONGL 311
 RESULT 4
 MPAL_CUPAR STANDARD; PRT; 346 AA.
 ID MPAL_CUPAR
 AC Q9SCG9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Major pollen allergen Cup a 1.
 OS Cupressus arizonica.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 OC NCBI_Taxid=49011;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=20571526; PubMed=1122214;
 RA Aceituno E., Del Pozo V., Minguez A., Arrieta I., Cortegano I.,
 RA Cardaba B., Gallardo S., Rojo M., Palomino P., Lahoz C.;

RT "Molecular cloning of major allergen from Cupressus arizonica pollen:
 RT Cup a 1." ;
 RL Clin. Exp. Allergy 30:1750-1758(2000).
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A II/AMB A II/CRY J I SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AJ243570; CAB62551.1; -;
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec lyase1.1.
 DR PRINTS; PRO0807; AMBALLERGEN.
 KM Allergen; Glycoprotein.
 FT CARBOHYD 127 127
 FT CARBOHYD 157 157
 FT CARBOHYD 272 272
 SQ SEQUENCE 346 AA; 37589 MW; F1281DCDA1D5DFD0 CRC64;
 Query Match 20.1%; Score 143; DB 1; Length 346;
 Best Local Similarity 53.8%; Pred. No. 1.2e-07;
 Matches 28; Conservative 8; Mismatches 16; Indels 0; Gaps 0;
 QY 73 RLMTIFSGNNITLKMMPYIAGYKTFDGRRAEVSYYVHNGAKFIRRVGDI 124
 DB 66 KALMTIFSGNNITLKMPLVYAGIKTIDGRGAVVHLGNGGCEPLMRASHT 117
 RESULT 5
 MPAL_CHAOR STANDARD; PRT; 375 AA.
 ID MPAL_CHAOR
 AC Q96385;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Major pollen allergen Cha o 1 precursor.
 OS Chamaecyparis obtusa (Japanese cypress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae;
 OC Chamaecyparis.
 OC NCBI_Taxid=13415;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RA TISSUE=Pollen;
 RX MEDLINE=96265194; PubMed=8676896;
 RA Suzuki M., Komiyama N., Itoh H., Sone T., Kuno K., Takagi I.,
 RA Ohta N.;
 RT "Purification, characterization and molecular cloning of Cha o 1, a
 RT major allergen of Chamaecyparis obtusa (Japanese cypress) pollen." ;
 RL Mol. Immunol. 33:451-460(1996).
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A II/AMB A II/CRY J I SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D45404; BAA08246.1; -;
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec lyase1.1.
 DR PRINTS; PRO0807; AMBALLERGEN.
 KM Allergen; Glycoprotein; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 375
 MAJOR POLLEN ALLERGEN CHA O 1.

FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 375 AA; 40258 MW; 81CD91DF7065DBB CRC64;

Query Match. 20.0%; Score 142; DB 1; Length 375;
 Best Local Similarity 57.7%; Pred. No. 1.7e-07;
 Matches 30; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

Qy 73 RELWITFGNNMNIKLKMPYIAGYTFDGRRAVSVYHNGAKFRVYDGI 124
 Db 87 RELWITFGNNMNIKLKMPYIAGYTFDGRRAVSVYHNGAKFRVYDGI 138

RESULT 6

9612_LYCES STANDARD; PRT; 404 AA.
 AC P24396;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Style development-specific protein 9612 precursor.
 GN 9612.

OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OC NCB1_TaxID=4081;

RP SEQUENCE FROM N.A.
 RC STRAIN=CV. VF36; TISSUE=Petiole;
 RX MEDLINE=91117185; PubMed=227637;
 RA Budelier K.A., Smith A.G., Gasser C.S.;
 RT "Regulation of a stylar transmitting tissue-specific gene in
 wild-type and transgenic tomato and tobacco."

RL Mol. Gen. Genet. 224:183-192(1990).
 CC -1- FUNCTION: MAY HAVE A ROLE IN THE DEVELOPMENT OF THE TRANSMITTING
 TISSUE OF THE STYLE AND/OR IN THE EVENTS RELATED TO POLLINATION
 SUCH AS SOME ASPECT IN THE FACILITATION OF COMPATIBLE POLLEN TUBE
 GROWTH.

CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY FOUND IN THE PISTIL WHERE IT IS
 FOUND IN THE OUTER FIVE LAYERS OF THE STRANDS OF TRANSMITTING
 TISSUE WITHIN THE UPPER TWO-THIRDS OF THE STYLE. FOUND AT MUCH
 LOWER LEVELS IN THE ANTHERS AND VEGETATIVE ORGANS.

CC -1- DEVELOPMENTAL STAGE: MAXIMUM LEVELS ARE FOUND DURING ANTHERESIS.
 CC -1- SIMILARITY: 54% IDENTICAL TO TOMATO PROTEINS P59 (AC P157122)
 AND P56 (AC P15721).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; X55193; CA38979.1; -;
 DR PIR; S12209; S12209.
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 KW Signal.

FT CHAIN 1 20 POTENTIAL.
 FT CARBOHYD 21 404 STYLE DEVELOPMENT-SPECIFIC PROTEIN 9612.
 FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 404 AA; 44298 MW; B26ED69B128D8675 CRC64;

Query Match 14.5%; Score 103; DB 1; Length 404;
 Best Local Similarity 30.8%; Pred. No. 0.0021;

Matches 33; Conservative 13; Mismatches 31; Indels 30; Gaps 4;

Qy 28 GRRIDFASKNFHLOKKTIGTRKWKNNRIW-----LQFAKLTGFTLM 70
 Db 67 GKSPAFSFCALGFGKMAIG-----GKNGRIYVTDGNDPVPKPGTLRAH-----VI 116

Qy 71 GRRPLWITFGNNMNIKLKMPYIAGYTFDGRRAVSVYHNGAKFI 117
 Db 117 ODEPLWITFGNNMNIKLKMPYIAGYTFDGRRAVSVYHNGAKFI 160

RESULT 7

WPI1_AMBAR
 ID WPI1_AMBAR STANDARD; PRT; 396 AA.

AC P27759;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pollen allergen Amb a 1.1 precursor (Antigen E) (Antigen Amb a 1).
 OS Ambrosia artemisiifolia (Short ragweed); Embryophyta; Tracheophyta;
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
 OC Heliantheae; Ambrosia.
 OC NCB1_TaxID=4212;

RN NCB1 [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC TISSUE=Pollen;
 RX MEDLINE=91093235; PubMed=1702434;
 RA Rafnar T., Griffith I.J., Kuo M.-C., Bond J.F., Rogers B.L.,
 RA Klapper D.G.;

RT "Cloning of Amb a I (antigen E), the major allergen family of short
 ragweed pollen."
 RL J. Biol. Chem. 266:1229-1236(1991).
 RN [2]

RP SEQUENCE FROM N.A., AND VARIANTS.

RC TISSUE=Pollen;

RX MEDLINE=92234570; PubMed=1809687;

RA Griffith I.J., Pollack J., Klapper D.G., Rogers B.L., Nault A.K.;
 RT "Sequence polymorphism of Amb a I and Amb a II, the major allergens
 in Ambrosia artemisiifolia (short ragweed)."

RL Int. Arch. Allergy Appl. Immunol. 96:296-304(1991).
 CC -1- SUBUNIT: MONOMER.

CC -1- TISSUE SPECIFICITY: POLLEN AND FLOWERS.

CC -1- PTM: The N-terminus is blocked.

CC -1- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.

CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 AMB A I/AMB A II/CRY J I SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; M63116; - NOT ANNOTATED_CDS.
 DR EMBL; M80558; AAA32665.1; -;
 DR PIR; A39099; A39099.
 DR PIR; A53240; A53240.
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 DR Antigen; Allergen; Signal; Multigene family; Polymorphism.
 KW Signal.

FT CHAIN 1 25 POTENTIAL.
 FT CARBOHYD 26 396 POLLEN ALLERGEN AMB A 1.1.
 FT VARIANT 92 92 E -> D.
 SQ SEQUENCE 396 AA; 42709 MW; 0CE7DDECB28841D CRC64;

Query Match 13.9%; Score 98.5; DB 1; Length 396;
 Best Local Similarity 34.7%; Pred. No. 0.006; Mismatches 22; Indels 15; Gaps 3;

QY 73 RPLMIIFSGNNIKLKMPLYAGYKTFDGRRAEVSYYH-----VNGAK--FIRAVD----- 121
 DB 113 RPLMIIFRDWYIRLBDKEMVYNSDKTIDGRGAKVEIINAGFTLNGVKNVITIHINIMEDVK 172
 QY 122 -----GIIAAYQNPAS 132
 DB 173 VNGGLIKNSDGPAA 187
 RESULT 8
 MP13_AMBAR STANDARD; PRT; 397 AA.
 ID MP13_AMBAR
 AC P27761;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pollen allergen Amb a 1.3 precursor (Antigen E) (Antigen Amb a I).
 OS Ambrosia artemisiifolia (Short ragweed).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
 OC Heliantheae; Ambrosia.
 OC NCBI_TaxID=4212;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pollen;
 RX MEDLINE=91093235; PubMed=1702434;
 RA Rafnar T., Griffith I.J., Kuo M.-C., Bond J.F., Rogers B.L.,
 RA Klapper D.G.;
 RT "Cloning of Amb a I (antigen E), the major allergen family of short
 RT ragweed pollen.";
 RL J. Biol. Chem. 266:1229-1236(1991).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANTS.
 RC TISSUE=Pollen;
 RX MEDLINE=92234570; PubMed=1809687;
 RA Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.;
 RT "Sequence polymorphism of Amb a I and Amb a II, the major allergens
 RT in Ambrosia artemisiifolia (short ragweed).";
 RL Int. Arch. Allergy Appl. Immunol. 96:296-304(1991).
 CC -1- SUBUNIT: MONOMER.
 CC -1- TISSUE SPECIFICITY: POLLEN AND FLOWERS.
 CC -1- PTM: The N-terminus is blocked.
 CC -1- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: M62961; AAA32668.1; -;
 CC EMBL: M80860; AAA32669.1; ALT_SEQ.
 CC PIR: C39099; C39099.
 CC InterPro: IPR002022; Amb allergen.
 CC Pfam: PF00544; pec_lyase; 1.
 CC PRINTS: PR00807; AMBALLERGEN.
 CC Antigen, Allergen, Signal, Multigene family, Polymorphism.
 FT SIGNAL 1 25
 FT CHAIN 26 397
 FT VARIANT 48 48 L -> Y.
 FT SEQUENCE 397 AA; 42928 MW; C8DB41257590DD0A CRC64;
 Query Match 13.6%; Score 97; DB 1; Length 397;
 Best Local Similarity 23.5%; Pred. No. 0.0086;
 Matches 31; Conservative 7; Mismatches 33; Indels 34; Gaps 3;
 QY 52 WNNRRIWQ-----FAKLT-----GTLNGRRRLMIIF 79

DB 61 WNNRQALDCAQGFAGKGYGKMGDVTYVTSNLDVYANPEKGTLLRPAQAQNRFLMIIF 120
 QY 80 SGNMNIKLMKMPYIAGYKTFDGRRAEVSYYHNGAKFIRVDGII 124
 DB 121 KNDVYINLNGELVNSDKTIDGRGVKVEI--INGSLTLMNKNII 163
 RESULT 9
 PEL_TOBAC STANDARD; PRT; 397 AA.
 ID PEL_TOBAC
 AC P40572;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Pectate lyase precursor (EC 4.2.2.2).
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OC NCBI_TaxID=4097;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Samsun; TISSUE=Pollen;
 RX MEDLINE=93043039; PubMed=1421152;
 RA Rogers H.J., Harvey A., Lonsdale D.M.;
 RT "Isolation and characterization of a tobacco gene with homology to
 RT pectate lyase which is specifically expressed during
 RT microsporogenesis.";
 RL Plant Mol. Biol. 20:493-502(1992).
 CC -1- CATALYTIC ACTIVITY: Eliminative cleavage of pectate to give
 CC oligosaccharides with 4-deoxy-alpha-D-gluc-4-enuronosyl groups at
 CC their non-reducing ends.
 CC -1- DEVELOPMENTAL STAGE: MAXIMALLY EXPRESSED LATE IN POLLEN
 CC DEVELOPMENT.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: X67158; CAA47630.1; -;
 CC EMBL: X67159; CAA47631.1; -;
 CC EMBL: X61102; CAA43414.1; -;
 CC PIR: S26211; S26211.
 CC InterPro: IPR002022; Amb allergen.
 CC Pfam: PF00544; pec_lyase; 1.
 CC PRINTS: PR00807; AMBALLERGEN.
 CC Lyase, Signal.
 FT SIGNAL 1 25
 FT CHAIN 26 397
 FT ACT SITE 272 272
 FT CARBOHYD 134 134
 FT CARBOHYD 227 227
 FT CONFLICT 156 156
 FT CONFLICT 189 189
 FT CONFLICT 200 200
 FT CONFLICT 202 202
 FT CONFLICT 249 249
 FT SEQUENCE 397 AA; 44351 MW; EFOA82CE5DA7643F CRC64;
 Query Match 13.1%; Score 93; DB 1; Length 397;
 Best Local Similarity 42.9%; Pred. No. 0.022;
 Matches 21; Conservative 7; Mismatches 17; Indels 4; Gaps 2;
 QY 67 FTLMGRRLMIIFSGNNIKLKMPLYAGYKTFDGRRAEVSYYH-NGA 114
 DB 105 YGVIOKEPLMIIFGKMKIKLSRELIVTSNKTIDGRGFN---VHIONGA 150

RESULT 10
MP12 AMBAR STANDARD; PRT; 398 AA.
ID MP12 AMBAR
AC P27760;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
OS Pollen allergen Amb a 1.2 precursor (Antigen E) (Antigen Amb a 1).
OC Ambrosia artemisiifolia (Short ragweed); Eudicotyledons; Tracheophyta;
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; easterids II; Asterales; Asteraceae; Asteroideae;
OC Heliantheae; Ambrosia.
OX NCBI_TaxID=4212;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pollen;
RX MEDLINE=91093235; PubMed=1702434;
RA Ratnar T., Griffith I.J., Kuo M.-C., Bond J.F., Rogers B.L.,
RT Klapper D.G.;
RT "Cloning of Amb a I (antigen E), the major allergen family of short
ragweed pollen."
RT J. Biol. Chem. 266:1229-1236 (1991).
[2]
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS.
RC TISSUE=Pollen;
RX MEDLINE=92234570; PubMed=1809687;
RA Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.;
RT "Sequence polymorphism of Amb a I and Amb a II, the major allergens
in Ambrosia artemisiifolia (short ragweed)."
RT Int. Arch. Allergy Appl. Immunol. 96:296-304 (1991).
RL Int. Arch. Allergy Appl. Immunol. 96:296-304 (1991).
CC -1- SUBUNIT: MONOMER.
CC -1- TISSUE SPECIFICITY: POLLEN AND FLOWERS.
CC -1- PTM: The N-terminus is blocked.
CC -1- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.
CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC AMB A I/AMB A II/CRY J I SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M62981; AAA32667.1; -;
CC EMBL: M60559; AAA32667.1; -;
CC PIR: B39099; B39099.
CC InterPro: IPR002022; Amb allergen.
CC Pfam: PF00544; Pec lyase; 1.
CC PRINTS: PR00807; AMBALLERGEN.
CC KMW Antigen; Allergen; Signal; Multigene family; Polymorphism.
CC FT SIGNAL 1 25
CC FT CHAIN 26 398 POTENTIAL.
CC FT VARIANT 345 345 POLLEN ALLERGEN AMB A 1.2.
CC FT VARIANT 381 381 R -> K.
CC FT VARIANT 381 381 L -> I.
CC SQ SEQUENCE 398 AA; 43664 MW; 020DC662D9B7416C CRC64;

Query Match 13.1%; Score 93; DB 1; Length 398;
Best Local Similarity 30.5%; Pred. No. 0.022;
Matches 32; Conservative 5; Mismatches 34; Indels 34; Gaps 3;

QY 52 WKNNRIWQ-----FAKLT-----GFTLMGRPPLMTIF 79
DB 62 WANNRQALADAGFACTGYGKRGDYVTYSDKDDVANPKSGTALFAAQRPLMTIF 121
QY 80 GGNNNIKIKMPMYIAGYKTFDGRRAEVSYYHNGAKFIRYVDGII 124
DB 122 KRNNAVHINQELVNVSDKTIDRGKVKVI--VNAGLTLMNVKNII 164

RESULT 11
MP12 AMBAR STANDARD; PRT; 397 AA.
ID MP12 AMBAR
AC P27762;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
OS Pollen allergen Amb a 2 precursor (Antigen K) (Antigen Amb a II).
OC Ambrosia artemisiifolia (Short ragweed).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; easterids II; Asterales; Asteraceae; Asteroideae;
OC Heliantheae; Ambrosia.
OX NCBI_TaxID=4212;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Flower;
RX MEDLINE=92013060; PubMed=1717566;
RA Rogers B.L., Moirgenstern J.P., Griffith I.J., Yu X.-B.,
RA Counsell C.M., Brauer A.W., King T.P., Gorman R.D., Kuo M.-C.C.;
RT "Complete sequence of the allergen Amb alpha II. Recombinant
RT expression and reactivity with T cells from ragweed allergic
RT patients."
RT J. Immunol. 147:2547-2552 (1991).
[2]
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS.
RC TISSUE=Flower;
RX MEDLINE=92234570; PubMed=1809687;
RA Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.;
RT "Sequence polymorphism of Amb a I and Amb a II, the major allergens
in Ambrosia artemisiifolia (short ragweed)."
RT Int. Arch. Allergy Appl. Immunol. 96:296-304 (1991).
RL Int. Arch. Allergy Appl. Immunol. 96:296-304 (1991).
CC -1- SUBUNIT: MONOMER.
CC -1- TISSUE SPECIFICITY: POLLEN AND FLOWERS.
CC -1- PTM: The N-terminus is blocked.
CC -1- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.
CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC AMB A I/AMB A II/CRY J I SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M80561; AAA32671.1; -;
CC PIR: A46469; A46469.
CC InterPro: IPR002022; Amb allergen.
CC Pfam: PF00544; Pec lyase; 1.
CC PRINTS: PR00807; AMBALLERGEN.
CC KMW Antigen; Allergen; Signal; Polymorphism.
CC FT SIGNAL 1 25
CC FT CHAIN 26 397 POTENTIAL.
CC FT VARIANT 70 70 N -> D (DETECTED ONLY IN FLOWER DNA).
CC FT VARIANT 138 138 K -> T (DETECTED ONLY IN FLOWER DNA).
CC FT VARIANT 321 321 K -> R (DETECTED ONLY IN FLOWER DNA).
CC SQ SEQUENCE 397 AA; 44082 MW; C78617E4C9A3D1DD CRC64;

Query Match 12.9%; Score 92; DB 1; Length 397;
Best Local Similarity 42.3%; Pred. No. 0.028;
Matches 22; Conservative 7; Mismatches 21; Indels 2; Gaps 1;

QY 73 RPLWTFSGNNNIKIKMPMYIAGYKTFDGRRAEVSYYHNGAKFIRYVDGII 124
DB 114 RPLWTFGRDMIIYLQGMNVVTSKTIIDGKAKVELVY--GGITLMNVKNVI 163
RESULT 12
ID P566_LYCES
AC P15721, STANDARD; PRT; 398 AA.

DT 01-APR-1990 (Rel. 14, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Probable pectate lyase P56 precursor (EC 4.2.2.2).
 GN LAT56.
 OS Lycopersicon esculentum (tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. VP36; TISSUE=Anther;
 RX MEDLINE=91322485; PubMed=1983191;
 RA Wing R.A., Yamaguchi J., Larabell S.K., Ursin V.M., McCormick S.;
 RT "Molecular and genetic characterization of two pollen-expressed genes
 that have sequence similarity to pectate lyases of the plant pathogen
 Erwinia.";
 RL Plant Mol. Biol. 14:17-28(1990).
 RL [2]
 RP REVISIONS.
 RC STRAIN=cv. VP36; TISSUE=Anther;
 RA Wing R.A.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MIGHT BE NEEDED DURING POLLEN DEVELOPMENT AND TUBE
 CC GROWTH.
 CC -1- CATALYTIC ACTIVITY: Eliminative cleavage of pectate to give
 CC oligosaccharides with 4-deoxy-alpha-D-gluc-4-enuronosyl groups at
 CC their non-reducing ends.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ANTHERS AND POLLEN.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; X15500; CAA3524.1; -
 CC PIR; S08383; S08383.
 DR InterPro: IPR002022; Amb_allergen.
 DR Pfam: PF00544; pec_lyase.1.
 DR PRINTS: PR00807; AMBALALGERN.
 KW Lyase; Multigene family; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 398 OR 22 (POTENTIAL).
 FT ACT SITE 273 273 PROBABLE PECTATE LYASE P56.
 FT CARBOHYD 135 135 POTENTIAL.
 FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 398 AA; 44563 MW; 8D676250BD8C7C8 CRC64;
 QY
 Query Match 12.8%; Score 91; DB 1; Length 398;
 Best Local Similarity 40.8%; Pred. No. 0.036;
 Matches 20; Conservative 10; Mismatches 15; Indels 4; Gaps 2;
 QY 67 FLNMRRLPIITFGNNMNIKMKPMYIAGKTPDGRRAVSVYVNGAFIRVDCII 114
 Db 106 FEVIOKGPLMITFARSMRIRLTRELIVSNKTIIDRG--KXVHLANGA 151
 RESULT 13
 MP14_AMBAR STANDARD; PRT; 392 AA.
 ID MP14_AMBAR
 AC P28744;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pollen allergen Amb a 1.4 precursor (Antigen E) (Antigen Amb a 1).
 DE Ambrosia artemisiifolia (Short ragweed).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
 OC Heliantheae; Ambrosia.
 OX NCBI_TaxID=4212;
 RN
 RP SEQUENCE FROM N.A., AND VARIANTS.
 RC TISSUE=Pollen;
 RX MEDLINE=92234570; PubMed=1809687;
 RA Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.;
 RT "Sequence polymorphism of Amb a 1 and Amb a 1L, the major allergens
 in Ambrosia artemisiifolia (short ragweed).";
 RL Int. Arch. Allergy Appl. Immunol. 96:296-304(1991).
 RL [1]
 CC -1- SUBUNIT: MONOMER.
 CC -1- TISSUE SPECIFICITY: POLLEN AND FLOWERS.
 CC -1- PTR: The N-terminus is blocked.
 CC -1- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A 1/AMB A 1I/CRY J I SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; M80562; AAA32670.1; -
 CC PIR; D53240; D53240.
 DR InterPro: IPR002022; Amb_allergen.
 DR Pfam: PF00544; pec_lyase.1.
 DR PRINTS: PR00807; AMBALALGERN.
 KW Antigen; Allergen; Signal; Multigene family; Polymorphism.
 FT SIGNAL 1 25
 FT CHAIN 26 392
 FT VAR1ANT 182 188 SHDGPV -> CNDGPV.
 SQ SEQUENCE 392 AA; 42842 MW; 7B6219C12P365DA9 CRC64;
 QY
 Query Match 12.5%; Score 89; DB 1; Length 392;
 Best Local Similarity 44.2%; Pred. No. 0.058;
 Matches 23; Conservative 7; Mismatches 20; Indels 2; Gaps 1;
 QY 73 RPLMIFSGNNMNIKMKPMYIAGKTPDGRRAVSVYVNGAFIRVDCII 124
 Db 114 RPLMIFRDMVIRLRELAINDKTIIDGKAVETINAGFA--IYVKNII 163
 RESULT 14
 ID PEL_LILLO STANDARD; PRT; 434 AA.
 AC P40973;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Pectate lyase precursor (EC 4.2.2.2).
 OS Lilium longiflorum (Trumpet lily).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae;
 OC Lilium.
 OX NCBI_TaxID=4690;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nellie white; TISSUE=Pollen;
 RA Kim S.R., Finkel D.J., An G.;
 RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Eliminative cleavage of pectate to give
 CC oligosaccharides with 4-deoxy-alpha-D-gluc-4-enuronosyl groups at
 CC their non-reducing ends.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: Z17328; CAA78976.1; -
 DR EMBL: L18911; AAA33398.1; -
 DR PIR: S29612; S29612.
 DR InterPro: IPR002022; Amb_allergen.
 DR Pfam: PF00544; pec_lyase.1
 DR PRINTS: PRO0507; AMBALERGEN.
 DR Lyase; Signal.
 KW SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 434 PEPTATE LYASE.
 FT ACT SITE 312 312 POTENTIAL.
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 434 AA; 48457 MW; C1F3E30AD2BBD064 CRC64;

Query Match 12.5%; Score 89; DB 1; Length 434;
 Best Local Similarity 55.9%; Pred. No. 0.064;
 Matches 19; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 73 RPLMTIFSGNNIKLKMPTIAGYKTFDGRRAV 106
 DB 151 RPLMTIFGKSNVIRLQELIINNKTIDGRANV 184

RESULT 15
 PES9_LYCES
 ID PES9_LYCES STANDARD; PRT; 449 AA.
 AC P15722;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Probable pectate lyase P59 precursor (EC 4.2.2.2).
 GN LATS9.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OC NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. VF36; TISSUE=anther;
 RX MEDLINE=91322485; PubMed=1983191;
 RA Wang R.A., Yamaguchi J., Larabell S.K., Ursin V.M., McCormick S.;
 RT "Molecular and genetic characterization of two pollen-expressed genes
 RT that have sequence similarity to pectate lyases of the plant pathogen
 RT Erwinia.";
 RL Plant Mol. Biol. 14:17-28(1990).
 CC -1- FUNCTION: MIGHT BE NEEDED DURING POLLEN DEVELOPMENT AND TUBE
 CC GROWTH.
 CC -1- CATALYTIC ACTIVITY: Eliminative cleavage of pectate to give
 CC oligosaccharides with 4-deoxy-alpha-D-gluc-4-enuronosyl groups at
 CC their non-reducing ends.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ANTHERS AND POLLEN.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation on
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: X15499; CAA33523.1; -
 DR PIR: S27098; S27098.
 DR InterPro: IPR002022; Amb_allergen.
 DR Pfam: PF00544; pec_lyase.1.
 DR PRINTS: PRO0807; AMBALERGEN.
 KW Lyase; Multigene family; Signal.
 FT SIGNAL 1 22 POTENTIAL.

FT CHAIN 23 449 PROBABLE PEPTATE LYASE P59.
 FT ACT SITE 325 325 POTENTIAL.
 FT CARBOHYD 56 56 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 449 AA; 50893 MW; 17E3AA13F173B03C CRC64;

Query Match 11.6%; Score 82.5; DB 1; Length 449;
 Best Local Similarity 22.3%; Pred. No. 0.32;
 Matches 39; Conservative 13; Mismatches 48; Indels 75; Gaps 7;

QY 3 VYVAFNPGPNRRVFIKXVSNVIHGRIDIFASKNFHLQKNTIGTGR----- 51
 DB 57 VYLAENQ--KLDTWKELKKV-----KGTINNSTRGDGTGKXKTCGCMVTP 101

QY 52 ---WNNRIRLQFAKLTGFTLMG----- 71
 DB 102 IDKWRCDENWADNRKKLADCAMGFSKALGGKDFYVVTDSVDYNDPRGTLRHAVI 161

QY 72 -RRPLMTIFSGNNIKLKMPTIAGYKTFDGRRAVSVVH-NGA---KFLIRV 120
 DB 162 QKEPLMTIFKRGKNIIRLHQEMTMSDKTIDARGV---VHITKGAGITLYIKNV 213

Search completed: April 23, 2003, 18:05:13
 Job time : 16 secs

GenCore version 5.1.4.p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 18:04:50 ; Search time 85 Seconds

(without alignments)
324.827 Million cell updates/sec

Title: US-09-142-524D-3

Perfect score: 711

Sequence: 1 MKYTVAFNQGPNRRVPIKR.....KFIKRVGIIAAYQNPASWK 134

Scoring table:

BLOSUM62

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP archaea: *
2: SP bacteria: *
3: SP fungi: *
4: SP human: *
5: SP invertebrate: *
6: SP mammal: *
7: SP mhc: *
8: SP organelle: *
9: SP phage: *
10: SP plant: *
11: SP rodent: *
12: SP virus: *
13: SP vertebrate: *
14: SP unclassified: *
15: SP xvirus: *
16: SP bacteriophage: *
17: SP archaea: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	188	26.4	374	10 Q8RUR1	Q8RUR1 cryptomeria
2	150	21.1	367	10 Q9LRT2	Q9LRT2 juniperus v
3	150	21.1	367	10 Q9LRT1	Q9LRT1 juniperus v
4	150	21.1	367	10 Q9M4S6	Q9M4S6 cupressus s
5	150	21.1	367	10 Q9M4S4	Q9M4S4 cupressus s
6	150	21.1	367	10 Q9M4S3	Q9M4S3 cupressus s
7	150	21.1	367	10 Q9M4S2	Q9M4S2 cupressus s
8	150	21.1	367	10 Q9M4S1	Q9M4S1 juniperus s
9	146	20.5	367	10 Q9M4S5	Q9M4S5 cupressus s
10	144	20.3	367	10 Q9M4S6	Q9M4S6 cupressus a
11	110.5	15.5	501	10 Q93Z04	Q93Z04 arabisidopsis
12	110.5	15.5	542	10 Q9SV40	Q9SV40 arabisidopsis
13	107.5	15.1	507	10 Q9FY19	Q9FY19 juniperus a
14	104.5	14.7	463	10 Q9SCP2	Q9SCP2 arabisidopsis
15	100.5	14.1	453	10 Q94LR5	Q94LR5 cryza sativ
16	99	13.9	472	10 Q9LFP5	Q9LFP5 arabisidopsis

17	97	13.6	455	10 Q64510	Q64510 arabisidopsis
18	95.5	13.4	181	10 Q23666	Q23666 arabisidopsis
19	95.5	13.4	410	10 Q9FMK5	Q9FMK5 arabisidopsis
20	95.5	13.4	432	10 Q93Z25	Q93Z25 arabisidopsis
21	95.5	13.4	459	10 Q23665	Q23665 arabisidopsis
22	95.5	13.4	459	10 Q9M552	Q9M552 arabisidopsis
23	95	13.4	392	10 Q9FM66	Q9FM66 arabisidopsis
24	95	13.4	452	10 Q9LRM5	Q9LRM5 arabisidopsis
25	94.5	13.3	226	10 Q23667	Q23667 arabisidopsis
26	94.5	13.3	354	10 Q94001	Q94001 arabisidopsis
27	94.5	13.3	398	10 Q9M505	Q9M505 vitis vinif
28	94.5	13.3	404	10 Q9S871	Q9S871 arabisidopsis
29	94.5	13.3	408	10 Q9C5M8	Q9C5M8 arabisidopsis
30	94.5	13.3	431	10 Q23017	Q23017 arabisidopsis
31	93.5	13.2	392	10 Q9FXD8	Q9FXD8 arabisidopsis
32	93.5	13.2	407	10 Q9SDM4	Q9SDM4 arabisidopsis
33	93.5	13.2	408	10 Q8W116	Q8W116 arabisidopsis
34	93	13.1	450	10 Q40319	Q40319 medicago sa
35	91.5	12.9	368	10 Q9C8G4	Q9C8G4 arabisidopsis
36	90.5	12.7	409	10 Q93XJ1	Q93XJ1 salix gligt
37	90	12.7	409	10 Q9LRT0	Q9LRT0 arabisidopsis
38	89.5	12.6	324	10 Q8S345	Q8S345 capsicum an
39	89.5	12.6	398	10 Q43783	Q43783 musa acumin
40	88.5	12.4	297	10 Q9FY87	Q9FY87 arabisidopsis
41	88.5	12.4	438	10 Q43862	Q43862 zea mays (m
42	87	12.2	368	10 Q94FT5	Q94FT5 fragaria an
43	85.5	12.0	416	10 Q9M828	Q9M828 arabisidopsis
44	85	12.0	227	10 Q23668	Q23668 arabisidopsis
45	84.5	11.9	401	10 Q24554	Q24554 zinnia eleg

ALIGNMENTS

RESULT 1

Q8RUR1 PRELIMINARY; PRT; 374 AA.

AC Q8RUR1; 01-JUN-2002 (TREMUR1, 21, Created)

DT 01-JUN-2002 (TREMUR1, 21, Last sequence update)

DT 01-JUN-2002 (TREMUR1, 21, Last annotation update)

DE Cry j 1 precursor.

GN Cry j 1.1 OR Cry j 1.2.

OS Cryptomeria japonica (Japanese cedar).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.

OK NCBI_TaxID=3369;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=POLLEN;

RA Futamura N., Shinohara K.;

RT "Isolation and characterization of cDNAs encoding major allergen Cry j

1 from Cryptomeria japonica pollen."

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB081309; BAB86286.1; -

DR EMBL; AB081310; BAB86287.1; -

KW Signal.

FT SIGNAL. 1 21 POTENTIAL.

FT CHAIN. 22 374 CRY j 1.

SO SEQUENCE 374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;

Query Match 26.4%; Score 188; DB 10; Length 374;
Best Local Similarity 71.2%; Pred. No. 2.6e-12;
Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 73 RPLMTISGNMNTIKLKMPTIAGYKTFDGRRAVSIVHWGAKFIRVDGI 124
Db 87 RPLMTISGNMNTIKLKMPTIAGYKTFDGRRAVSIVHWGAKFIRVDGI 134

RESULT 2

Q9LRT2 PRELIMINARY; PRT; 367 AA.

```

AC 09LTL2:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Pollen major allergen 1-2.
OS Juniperus virginiana (Eastern red cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
NCBI_TaxID=33584;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21315424; PubMed=11422137;
RA Midoro-Horicuti T.M., Goldblum R.M., Brooks E.G.;
RT "Identification of mutations in the genes for the pollen allergens of
RT eastern red cedar (Juniperus virginiana).";
RL Clin. Exp. Allergy 31:771-778(2001).
DR EMBL; AF151427; AAF80164.1;-.
DR InterPro; IPR020222; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39768 MW; 0A6AC2F1BAF89586 CRC64;

Query Match 21.1%; Score 150; DB 10; Length 367;
Best Local Similarity 55.8%; Pred. No. 3.7e-08;
Matches 29; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

OY 73 RPLMIIFSGNMNITKLMPIYAGKTFDGRRAEVSYHVNGAKFIRRYDGI 124
DB 87 KALMIIFSGNMNITKLMPIYAGKTFDGRGADVHLGNGPCLFMRKXSHV 138

RESULT 3
O9LTL1 PRELIMINARY; PRT; 367 AA.
AC O9LTL1:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Pollen major allergen 1-1.
OS Juniperus virginiana (Eastern red cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
NCBI_TaxID=33584;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21315424; PubMed=11422137;
RA Midoro-Horicuti T.M., Goldblum R.M., Brooks E.G.;
RT "Identification of mutations in the genes for the pollen allergens of
RT eastern red cedar (Juniperus virginiana).";
RL Clin. Exp. Allergy 31:771-778(2001).
DR EMBL; AF151429; AAF80166.1;-.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39708 MW; DCBD1981A74E4711 CRC64;

Query Match 21.1%; Score 150; DB 10; Length 367;
Best Local Similarity 55.8%; Pred. No. 3.7e-08;
Matches 29; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

OY 73 RPLMIIFSGNMNITKLMPIYAGKTFDGRRAEVSYHVNGAKFIRRYDGI 124
DB 87 KTLMIIFSGNMNITKLMPIYAGKTFDGRGADVHLGNGPCLFMRKXSHV 138

RESULT 4
O9M4S6 PRELIMINARY; PRT; 367 AA.
AC O9M4S6:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
NCBI_TaxID=13469;

```

```

GN CUPs1.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
NCBI_TaxID=13469;
RN [1]
RP SEQUENCE FROM N.A.
RX Monasale R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
RT of Cupressus sempervirens.";
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF257491; AAF72625.1;-.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39835 MW; B4E9C60108C2C5A3 CRC64;

Query Match 21.1%; Score 150; DB 10; Length 367;
Best Local Similarity 55.8%; Pred. No. 3.7e-08;
Matches 29; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

OY 73 RPLMIIFSGNMNITKLMPIYAGKTFDGRRAEVSYHVNGAKFIRRYDGI 124
DB 87 KALMIIFSGNMNITKLMPIYAGKTFDGRGADVHLGNGPCLFMRKXSHV 138

RESULT 5
O9M4S4 PRELIMINARY; PRT; 367 AA.
AC O9M4S4:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
GN CUPs1.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
NCBI_TaxID=13469;
RN [1]
RP SEQUENCE FROM N.A.
RX Monasale R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
RT of Cupressus sempervirens.";
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF257493; AAF72627.1;-.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39934 MW; 974D3011D74E3D6E CRC64;

Query Match 21.1%; Score 150; DB 10; Length 367;
Best Local Similarity 55.8%; Pred. No. 3.7e-08;
Matches 29; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

OY 73 RPLMIIFSGNMNITKLMPIYAGKTFDGRRAEVSYHVNGAKFIRRYDGI 124
DB 87 KALMIIFSGNMNITKLMPIYAGKTFDGRGADVHLGNGPCLFMRKXSHV 138

RESULT 6
O9M4S3 PRELIMINARY; PRT; 367 AA.
AC O9M4S3:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
GN CUPs1.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
NCBI_TaxID=13469;

```

```

DR Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ293767; CAC48400.1; -.
DR InterPro: IPR002022; Amb_allergen.
DR Pfam: PF00544; pec_lyase; 1.
SQ SEQUENCE 367 AA; 39808 MW; 5D28204DBFDB9D7 CRC64;

Query Match 21.1%; Score 150; DB 10; Length 367;
Best Local Similarity 55.8%; Pred. No. 3.7e-08;
Matches 29; Conservative 9; Mismatches 14; Indels 0; Gaps 0

QY 73 RPLMTIFSGNNIKLKMPLYAGYKTFDGRRAEVSVYHVNCAKFTIRVDGII 124
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 87 KALMTIFSGNNIKLKMPLYAGHKTTIDRGADVHLGNGGFCFLFMKRVSHVI 138

RESULT 9
Q9M4S5 PRELIMINARY; PRT; 367 AA.
ID Q9M4S5
AC Q9M4S5;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
GN CupS1.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX NCBI_TaxId=13469;
RN [1]
RP SEQUENCE FROM N.A.
RA Monsalve R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
RT of Cupressus sempervirens."
RT Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF257492; AAF72626.1; -.
DR InterPro: IPR002022; Amb_allergen.
DR Pfam: PF00544; pec_lyase; 1.
DR PRINTS: PR00807; AMBALBERGEN.
SQ SEQUENCE 367 AA; 39894 MW; 5D56FC0B3263B741 CRC64;

Query Match 20.5%; Score 146; DB 10; Length 367;
Best Local Similarity 53.8%; Pred. No. 1e-07;
Matches 28; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 73 RPLMTIFSGNNIKLKMPLYAGYKTFDGRRAEVSVYHVNCAKFTIRVDGII 124
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 87 KALMTIFSGNNIKLKMPLYAGHKTTIDRGADVHLGNGGFCFLFMKRVSHVI 138

RESULT 10
Q93XL6 PRELIMINARY; PRT; 367 AA.
ID Q93XL6
AC Q93XL6;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative allergen Cup a 1 precursor.
GN Cup A 1.
OS Cupressus arizonica.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX NCBI_TaxId=49011;
RN [1]
RP SEQUENCE FROM N.A.
RA Butteroni C., Di Felice G., Pini C.;
RT "Cloning of Cupressus arizonica major allergen."
RT Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ278498; CAC37790.2; -.
DR Signal.
KW SIGNAL.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 222 367 PUTATIVE ALLERGEN CUP A 1.
SQ SEQUENCE 367 AA; 39898 MW; 5D56FC0B3263B741 CRC64;

```

Db 237 IDIFASKRFEIEKCTIGT-----DDCVAVG TGSSNITIKDL

Db 237 IDIFASKRFEIEKCTIGT-----DDCVAVG TGSSNITIKDL

QY 91 MYAGYKTFD-----RAEVSYVHNGAKFIRVDGI 123
 DB 274 TCGPHGMSIGLKGKNSRSEVSFVHLDGAKFIDTQNGI 312

RESULT 14

09SCP2 PRELIMINARY; PRT; 463 AA.
 ID 09SCP2;
 AC 09SCP2;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Pectate lyase-like protein.
 GN T4D2.120.

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Nyakatura G., Fartmann B., Dauner D., Sterr W., Holland R.,
 RA Weichselgartner M., Mewes H.W., Lemcke K., Mayer K.F.X., Quettier F.,
 RA Salanoubat M.;
 RL Submitted (NCV-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]

RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL132958; CAB64222.1;
 DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALERGEN.
 KW Lyase.

SO SEQUENCE 463 AA; 51225 MW; 3424B0DE4ADBA604 CRC64;

Query Match 14.7%; Score 104.5; DB 10; Length 463;
 Best Local Similarity 30.5%; Pred. No. 0.0047;
 Matches 32; Conservative 8; Mismatches 32; Indels 33; Gaps 3;

QY 52 WKNNRIWL-----QPAKLTLG-----FTLMGRRLPMTIP 79
 DB 66 WPNNRQGLADCGIGFGYALGKGQGYFTDSSDDANVPKFTLRVGIQEBPLMTIV 125

QY 80 SGNNNITLKMPTIAGYKTFDGRRAEVSYVHNGAKFIRVDGII 124
 DB 126 PSNNMIRLKOELIFNSYKTLDRGANVHIVG-GGCITLQYVSNII 169

RESULT 15

094IRS PRELIMINARY; PRT; 453 AA.

DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Putative pectate lyase.

OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriocaridaceae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]

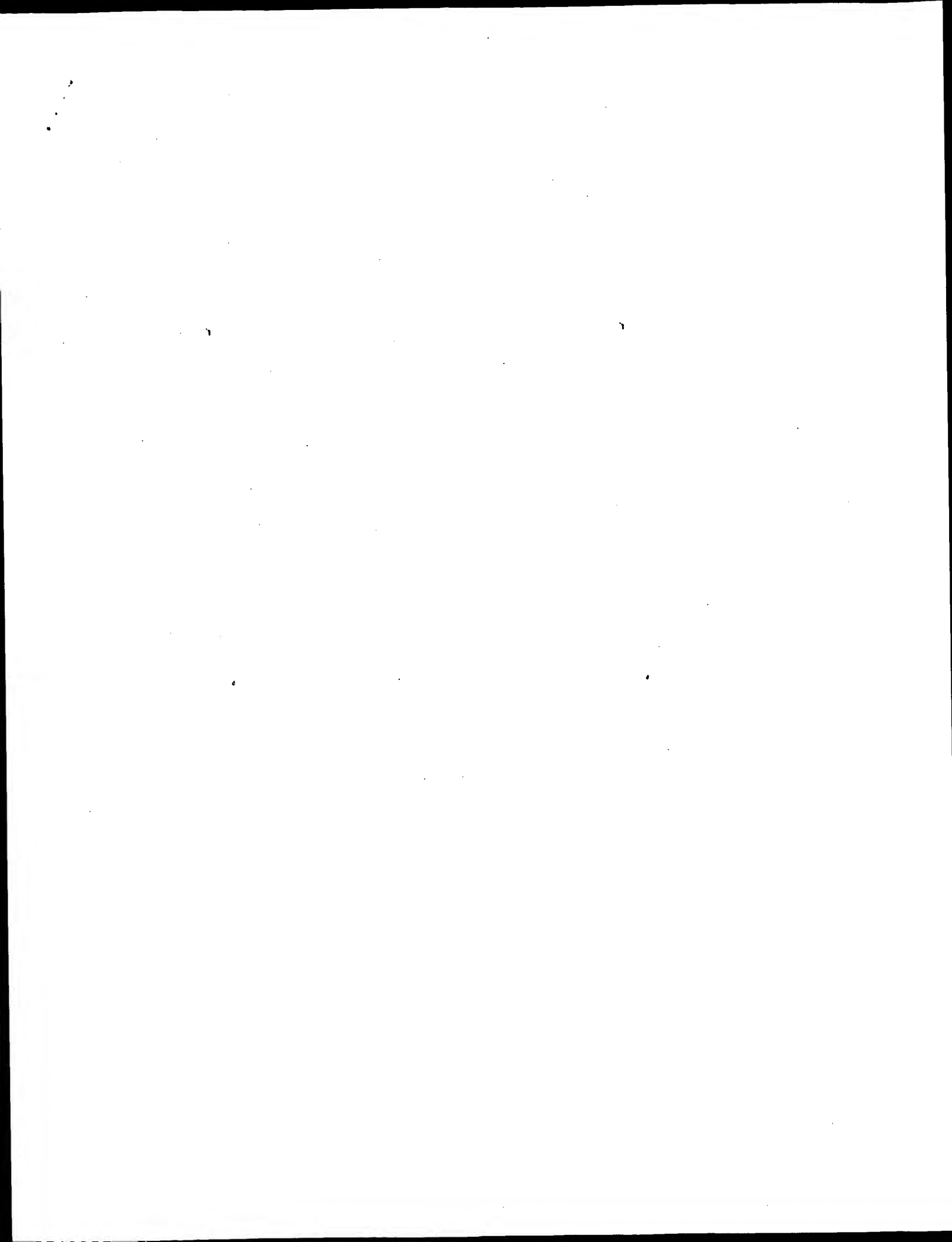
RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,
 RA Brenner M., Burgess S., Hance M., Shvartsbeyn M., Tslerin T.,
 RA Riggall F., Hsiao J., Zisman V., Blunt S., Pai G., Vanaken S.B.,
 RA Uteback T.R., Feldblum T.V., Quackenbush J., Salzberg S.L.,
 RA White O., Fraser C.M.;
 RT "Oryza sativa chromosome 10 BAC OSUNB0011A08 genomic sequence.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC034258; AAK54283.1; ..

DR InterPro; IPR002022; Amb_allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 KW Lyase.
 SO SEQUENCE 453 AA; 48006 MW; 1411BBE1A40901DC CRC64;

Query Match 14.1%; Score 100.5; DB 10; Length 453;
 Best Local Similarity 48.8%; Pred. No. 0.013;
 Matches 20; Conservative 7; Mismatches 13; Indels 1; Gaps 1;

QY 74 PLMIIFSGNNNITLKMPTIAGYKTFDGRRAEVSYVHNGA 114
 DB 140 PLMIIVRAGDNTIRLNEBLVNSYKTLDRGANV-HVGAGCA 179

Search completed: April 23, 2003, 18:07:08
 Job time : 89 secs



GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 18:04:50 ; Search time 36 Seconds

(without alignments)
495.988 Million cell updates/sec

Title: US-09-142-524D-3

Perfect score: 711

Sequence: 1 MKVTVAFNQSGNRRVFIKR.....KFTIRVDGIIAYGNPMSWK 134

Scoring table:

BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

```

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	711	100.0	134	18	AAW27371	Multi-epitope pep
2	523.5	73.6	105	18	AAW27370	Multi-epitope pep
3	320	45.0	80	18	AAW27369	Multi-epitope pep
4	198.5	27.9	214	22	AAW69120	Cedar pollen aller
5	188	26.4	353	16	AAW73388	Cedar pollen aller
6	188	26.4	353	17	AAW81587	Japanese cedar pol
7	188	26.4	373	20	AAW25664	Cedar pollen aller
8	188	26.4	374	20	AAW25668	Cedar pollen aller
9	188	26.4	373	14	AAW31937	Japanese cedar all
10	188	26.4	374	15	AAW45541	Cry j I. Cryptome
						Cry j I pollen all

11	188	26.4	374	15	AAW60166	Japanese cedar pol
12	188	26.4	374	16	AAW82490	Cry j I Japanese C
13	188	26.4	374	20	AAW25665	Cedar allergen 493
14	188	26.4	374	20	AAW25669	Japanese cedar all
15	178	25.0	210	22	AAW69103	Cedar pollen aller
16	162	22.8	96	22	AAW69104	Cedar pollen aller
17	152	22.8	96	22	AAW69109	Cedar pollen aller
18	152	22.8	97	22	AAW69105	Cedar pollen aller
19	150	21.1	367	15	AAW45577	Jun s I. Juniperu
20	150	21.1	370	15	AAW45578	Jun v I. Juniperu
21	148.5	20.9	460	16	AAW69791	Japonicum allergen
22	148.5	20.9	514	15	AAW53690	Japanese cedar pol
23	148.5	20.9	514	16	AAW74333	Japanese cedar pol
24	148.5	20.9	514	16	AAW69792	Japonicum allergen
25	148.5	20.9	514	17	AAW93599	Japan cedar pollen
26	148.5	20.9	514	17	AAW81586	Cedar pollen aller
27	148.5	20.9	514	20	AAW25666	Cedar pollen aller
28	148.5	20.9	514	20	AAW25667	Japanese cedar all
29	148	20.8	186	22	AAW69102	Japanese cedar all
30	142	20.0	354	17	AAW04344	Cedar pollen aller
31	142	20.0	354	19	AAW42121	Chamaecyparis obtu
32	142	20.0	375	17	AAW04345	Chamaecyparis obtu
33	141.5	19.9	93	22	AAW69117	Artificial sequenc
34	136	19.1	30	19	AAW44682	Artificial sequenc
35	120.5	16.9	74	21	AAW23905	Artificial sequenc
36	120.5	16.9	94	21	AAW23906	Artificial sequenc
37	120.5	16.9	95	21	AAW23907	Artificial sequenc
38	120.5	16.9	95	21	AAW23901	Artificial sequenc
39	120.5	16.9	95	22	AAW69098	Artificial sequenc
40	120.5	16.9	99	21	AAW23902	Cedar pollen aller
41	120.5	16.9	99	21	AAW23907	Artificial sequenc
42	120.5	16.9	99	22	AAW69099	Artificial sequenc
43	118.5	16.7	514	17	AAW04346	Cedar pollen aller
44	118.5	16.7	514	19	AAW42122	Chamaecyparis obtu
45	117	16.5	47	19	AAW60353	Japanese cypress p

ALIGNMENTS

```

RESULT 1
AAW27371
ID AAW27371 standard; peptide: 134 AA.
XX
AC AAW27371;
XX
DE 24-MAR-1998 (first entry)
XX
DE Multi-epitope peptide used as immunotherapeutic agent #3.
XX
KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;
KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
XX
OS Synthetic.
XX
PN WO9732600-A1.
XX
PD 12-SEP-1997.
XX
PF 10-MAR-1997; 97WO-JP00740.
XX
PR 10-MAR-1996; 96JP-0080702.
XX
PA (MEIP) MEIJI MILK PROD CO LTD.
XX
DA Dairiki K, Iwama A, Kino K, Kume A, Sone T;
XX
WP 1997-470495/43.
XX
PT Peptide immuno:therapeutic agent to treat allergic diseases -
PT contains multi-epitope peptide containing T cell epitope regions
PT from different allergens
XX

```

PS Claim 6; Page 32; 58pp; Japanese.
 XX The present sequence represents a multi-epitope peptide which is used as
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 CC or more different allergens (preferably linked via arginine or lysine
 CC dimers), where the T cell epitope regions have a positivity index
 CC greater than 100 as measured in a patient group responding to the
 CC allergen; have at least 70% reactivity with lymphocytes from patients
 CC responding to the allergen; and are not reactive with immunoglobulin E
 CC (IgE) antibodies from patients responsive to the allergen. The agent can
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
 XX Sequence 134 AA;
 SQ
 Query Match 100.0%; Score 711; DB 18; Length 134;
 Best Local Similarity 100.0%; Pred. No. 1,7e-82;
 Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKVTVAENQEGPNRRVFIKRVSNVTHGRRIDIFASKNFHLQKNTTGTGRMKNNRIWQ 60
 DB 1 FAKLTGFTLMGRRLMTIFSGNNMNIKLMPMYIAGYKTFDGRRAEVSYYHNGAKFIRRY 120
 QY 61 FAKLTGFTLMGRRLMTIFSGNNMNIKLMPMYIAGYKTFDGRRAEVSYYHNGAKFIRRY 120
 DB 61 FAKLTGFTLMGRRLMTIFSGNNMNIKLMPMYIAGYKTFDGRRAEVSYYHNGAKFIRRY 120
 QY 121 DGIITAAVQNPASWK 134
 DB 121 DGIITAAVQNPASWK 134
 Db 121 DGIITAAVQNPASWK 134
 RESULT 2
 AAM27370
 ID AAM27370 standard; peptide; 105 AA.
 XX AAM27370;
 AC
 XX 24-MAR-1998 (first entry)
 DT
 XX Multi-epitope peptide used as immunotherapeutic agent #2.
 DE
 XX Multi-epitope peptide; immunotherapeutic agent; allergic disease;
 KM T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
 XX
 OS Synthetic.
 XX WO9732600-A1.
 PN
 XX 12-SEP-1997.
 PD
 XX 10-MAR-1997; 97WO-JP00740.
 PF
 XX 10-MAR-1996; 96JP-0080702.
 PR
 XX (MEIP) MEIJI MILK PROD CO LTD.
 PA
 XX Dai-ichi K. Iwama A. Kino K. Kume A. Sone T;
 PI
 XX WPI; 1997-470495/43.
 DR
 XX Peptide immuno:therapeutic agent to treat allergic diseases -
 PT contains multi-epitope peptide containing T cell epitope regions
 PT from different allergens
 PT
 PS Claim 6; Page 31; 58pp; Japanese.
 XX The present sequence represents a multi-epitope peptide which is used as
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 CC or more different allergens (preferably linked via arginine or lysine
 CC dimers), where the T cell epitope regions have a positivity index
 CC greater than 100 as measured in a patient group responding to the
 CC allergen; have at least 70% reactivity with lymphocytes from patients
 CC responding to the allergen; and are not reactive with immunoglobulin E
 CC (IgE) antibodies from patients responsive to the allergen. The agent can
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
 XX Sequence 134 AA;
 SQ

CC responding to the allergen; and are not reactive with immunoglobulin E
 CC (IgE) antibodies from patients responsive to the allergen. The agent can
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
 XX Sequence 105 AA;
 SQ
 Query Match 73.6%; Score 523.5; DB 18; Length 105;
 Best Local Similarity 78.4%; Pred. No. 9.8e-59;
 Matches 105; Conservative 0; Mismatches 0; Indels 29; Gaps 2;
 QY 1 MKVTVAENQEGPNRRVFIKRVSNVTHGRRIDIFASKNFHLQKNTTGTGRMKNNRIWQ 60
 DB 1 MKVTVAENQEGPNRRVFIKRVSNVTHGRRIDIFASKNFHLQKNTTGTGRMKNNRIWQ 60
 QY 61 FAKLTGFTLMGRRLMTIFSGNNMNIKLMPMYIAGYKTFDGRRAEVSYYHNGAKFIRRY 120
 DB 61 FAKLTGFTLMGRRLMTIFSGNNMNIKLMPMYIAGYKTFDGRRAEVSYYHNGAKFIRRY 120
 QY 121 DGIITAAVQNPASWK 134
 DB 92 DGIITAAVQNPASWK 105
 Db 92 DGIITAAVQNPASWK 105
 RESULT 3
 AAM27369
 ID AAM27369 standard; peptide; 80 AA.
 XX AAM27369;
 AC
 XX 24-MAR-1998 (first entry)
 DT
 XX Multi-epitope peptide used as immunotherapeutic agent #1.
 DE
 XX Multi-epitope peptide; immunotherapeutic agent; allergic disease;
 KM T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
 XX
 OS Synthetic.
 XX WO9732600-A1.
 PN
 XX 12-SEP-1997.
 PD
 XX 10-MAR-1997; 97WO-JP00740.
 PF
 XX 10-MAR-1996; 96JP-0080702.
 PR
 XX (MEIP) MEIJI MILK PROD CO LTD.
 PA
 XX Dai-ichi K. Iwama A. Kino K. Kume A. Sone T;
 PI
 XX WPI; 1997-470495/43.
 DR
 XX Peptide immuno:therapeutic agent to treat allergic diseases -
 PT contains multi-epitope peptide containing T cell epitope regions
 PT from different allergens
 PT
 PS Claim 6; Page 31; 58pp; Japanese.
 XX The present sequence represents a multi-epitope peptide which is used as
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 CC or more different allergens (preferably linked via arginine or lysine
 CC dimers), where the T cell epitope regions have a positivity index
 CC greater than 100 as measured in a patient group responding to the
 CC allergen; have at least 70% reactivity with lymphocytes from patients
 CC responding to the allergen; and are not reactive with immunoglobulin E
 CC (IgE) antibodies from patients responsive to the allergen. The agent can
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
 XX Sequence 80 AA;
 SQ
 Query Match 45.0%; Score 320; DB 18; Length 80;

```

Qy      92 YIAGKTFDGRRAEVSYVHWGAKFIRRVGGIIAATONPASK 134
Db      113 -----DPR-----GIITAIOTNPASK 128

RESULT 5
AAR75388
ID AAR75388 standard; protein; 353 AA.
XX
AC AAR75388;
XX
DT 12-MAR-1996 (first entry)
XX
DE Japanese cedar pollen allergen Cryj I.
XX
KW Japanese cedar; pollen allergen; Cryj I; T-cell epitope; peptides;
KM prevention; treatment; cryptomeria pollinosis.
XX
OS Cryptomeria japonica.
XX
FH Key
FH Peptide
FT /note= "T-cell epitope peptide"
FT /note= "T-cell epitope peptide"
FT /note= "T-cell epitope peptide"
FT /note= "T-cell epitope peptide"
FT /note= "T-cell epitope peptide"
FT /note= "T-cell epitope peptide"
FT /note= "T-cell epitope peptide"
FT /note= "T-cell epitope peptide"
FT /note= "T-cell epitope peptide"
FT /note= "T-cell epitope peptide"
FT /note= "T-cell epitope peptide"
FT /note= "T-cell epitope peptide"
FT /note= "T-cell epitope peptide"
PN JP07118295-A.
PD 09-MAY-1995.
PF 20-OCT-1993; 93JP-0262626.
PR 20-OCT-1993; 93JP-0262626.
XX
PA (MEIP ) MEITI MILK PROD CO LTD.
XX
DR WPI; 1995-203834/27.
XX
PT New cryptomeria pollen allergen T-cell epitope peptide - used for
PT prevention, treatment and investigation of Japanese cedar pollinosis
XX
PS Disclosure; Figs 1-2; bpp; Japanese.
CC AAR75388 is the Japanese cedar pollen allergen Cryj I, from which the
CC T-cell epitope peptides AAR89289-R89295 were derived. The peptides
CC can be used for the prevention and treatment of cryptomeria
CC pollinosis, and also for the investigation of pollinosis.
XX
SO Sequence 353 AA;

Query Match 26.4%; Score 188; DB 16; Length 353;
Best Local Similarity 71.2%; Pred. No. 3.2e-15;
Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0

Oy      73 RPLWIFSGNMNKLKMPYINGTYTPFGRAEVSYVHWGAKFIRRVDDGI 124
Db      66 RPLWIFSGNMNKLKMPYINGTYTPFGRAEQVYIGNGPCVFIFKRVSNI 117

```

```

XX AC AAR81587;
XX XX
XX DT 24-MAY-1996 (first entry)
XX DE Cedar pollen allergen B.
XX KW Cedar; pollen; allergen; immunoglobulin E; IgE; T-cell epitope;
XX antibody; pollinosis; therapy; immunotherapy.
XX OS Cryptomeria japonica.
XX PN EP700929-A2.
XX PD 13-MAR-1996.
XX PF 08-SEP-1995; 95SEP-0306295.
XX PR 14-JUL-1995; 95JUL-0200221.
XX PR 10-SEP-1994; 94SEP-0242137.
XX PR 14-JUL-1995; 95JUL-0200204.
XX PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX PI Hino K, Saito S, Taniguchi Y;
XX DR WPI; 1996-140976/15.
XX PT New peptide(s) derived from cedar pollen allergens - activate
XX PT allergen-specific T-cells, but not allergen-specific IgE antibodies,
XX PT used for treating cedar pollinosis
XX PS Claim 5; Page 31-32; 36pp; English.
XX CC Synthetic peptides based on portions of cedar pollen allergens A
XX CC (AAR81586) and B (AAR81587) were tested for their ability to activate
XX CC cedar allergen-specific T-cells, but not allergen-specific IgE
XX CC antibodies. 6 Peptides (AAR81580-R81585) were identified as T-cell
XX CC epitopes. These peptides, plus subsequences (AAR81573-79) essential
XX CC for T-cell recognition, and homologous peptides (AAR81588-96) can
XX CC be used as immunotherapeutic agents to treat or prevent cedar
XX CC pollinosis, avoiding side-effects such as anaphylaxis.
XX SQ Sequence 353 AA;
XX
XX Query Match 26.4%; Score 188; DB 17; Length 353;
XX Best Local Similarity 71.2%; Pred. No. 3.2e-15;
XX Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
QY 73 RPLWTFSGNNMNIKMKMPYTAGYKTPDGRRAEVSYYHVNGAKFIRRVDDGI 124
DB 66 RPLWTFSGNNMNIKMKMPYTAGYKTPDGRGAQYIYGNGGCVFIKRVSNVI 117

RESULT 7
AAY25664
ID AAY25664 standard; protein; 373 AA.
XX AC AAY25664;
XX DT 30-SEP-1999 (first entry)
XX DE Cedar allergen 493634 Cry j IB protein fragment.
XX KW Major histocompatibility complex; class II; desensitizing; human;
XX KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
XX KW chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
XX KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
XX KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
XX KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
XX OS Cedrus sp.
XX XX

```

```

PN MN W09934826-A1.
XX XX
XX PD 15-JUL-1999.
XX XX
XX DT 11-JAN-1999; 99WO-GB00080.
XX PF 21-SEP-1998; 98GB-0020474.
XX PR 09-JAN-1998; 98GB-0000445.
XX PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX PI Kay AB, Larche M;
XX DR WPI; 1999-458255/38.
XX PT Desensitizing patients to polypeptide allergens
XX PS Example 6, Page 73; 117pp; English.
XX CC This invention describes a novel method of desensitizing a patient to a
XX CC polypeptide allergen and comprises administering to the patient a peptide
XX CC derived from the allergen where restriction to a MHC Class II molecule
XX CC possessed by the patient can be demonstrated for the peptide and the
XX CC peptide is able to induce a late phase response in an individual who
XX CC possesses the MHC Class II molecule. The methods can be used for
XX CC desensitizing patients to allergens present in e.g. grass, tree and weed
XX CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
XX CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
XX CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
XX CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
XX CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
XX CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
XX CC produce immunological vaccines which may be used to prevent and/or treat
XX CC conditions involving hypersensitivity to allergens. This sequence
XX CC represents a cedar (Cedrus sp.) allergen 493634 Cry j IB.
XX SQ Sequence 373 AA;
XX
XX Query Match 26.4%; Score 188; DB 20; Length 373;
XX Best Local Similarity 71.2%; Pred. No. 3.5e-15;
XX Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
QY 73 RPLWTFSGNNMNIKMKMPYTAGYKTPDGRRAEVSYYHVNGAKFIRRVDDGI 124
DB 86 RPLWTFSGNNMNIKMKMPYTAGYKTPDGRGAQYIYGNGGCVFIKRVSNVI 137

RESULT 8
AAY25668
ID AAY25668 standard; protein; 373 AA.
XX AC AAY25668;
XX DT 30-SEP-1999 (first entry)
XX DE Japanese cedar allergen 541803 Cry j I precursor protein fragment.
XX KW Major histocompatibility complex; class II; desensitizing; human;
XX KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
XX KW chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
XX KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
XX KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
XX KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
XX OS Cedrus sp.
XX PN W09934826-A1.
XX PD 15-JUL-1999.
XX PF 11-JAN-1999; 99WO-GB00080.
XX PR 21-SEP-1998; 98GB-0020474.

```

PR 09-JAN-1998; 98GB-0000445.
 XX
 PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 XX
 PI Kay AB, Larche M;
 XX
 DR WPI; 1999-458255/38.
 XX
 PT Desensitizing patients to polypeptide allergens
 XX
 PS Example 6; Page 75; 117pp; English.
 XX
 CC This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The methods can be used for
 CC desensitizing patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC Terribio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents a Japanese cedar (Cedrus sp.) allergen 541803 Cry j I
 CC precursor.
 CC
 SQ Sequence 373 AA;
 Query Match 26.4%; Score 188; DB 20; Length 373;
 Best Local Similarity 71.2%; Pred. No. 3.5e-15;
 Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
 QY 73 RPLWIIIFSGNNNITKIKMPYIAGYKTFDGRRAEVSYYHNGAKFIRRVDCII 124
 DB 86 RPLWIIIFSGNNNITKIKMPYIAGYKTFDGRGAQYVYIGNGCPVFIRKVSNI 137

RESULT 9
 AAR31937
 ID AAR31937 standard; Protein; 374 AA.
 AC AAR31937;
 XX
 DT 03-JUN-1993 (first entry)
 XX
 DE Cry j I.
 XX
 KM Japanese cedar pollen; allergen; antigen; allergy; B cell; T cell.
 XX
 OS Cryptomeria japonica.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..21
 FT /note= "signal peptide"
 FT 22..374
 FT Protein /note= "mature Cry j I"
 XX
 PN WO9301213-A.
 XX
 PD 21-JAN-1993.
 XX
 PF 10-JUL-1992; 92WO-US05661.
 XX
 PR 12-JUL-1991; 91US-0729134.
 PR 15-JUL-1991; 91US-0730452.
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 XX
 PI Bond JF, Griffith IU, Pollock J;

XX
 DR WPI; 1993-045434/05.
 DR N-PSDB; AAQ5304.
 XX
 PT Nucleic acid sequence encoding Cryptomeria japonica allergen
 PT for the diagnosis treatment and prevention of allergic reactions
 PT to Japanese cedar pollen
 XX
 PS Claim 11; Page 42; 69pp; English.
 XX
 CC Fresh pollen and stimulant cone samples were collected from a single
 CC Cryptomeria japonica (Japanese cedar) tree. RNA was prep'd. and used
 CC to synthesize cDNA. The cDNA was subjected to successive rounds of
 CC PCR to yield a full length Cry j I clone. Cry j I or an antigenic
 CC fragment of it may be used for detecting, treating and preventing an
 CC allergic response to Japanese cedar pollen allergen. It is capable of
 CC modifying both the B and T cell response to Cry j I and T cell response
 CC to a Cry j I antigen.
 CC
 SQ Sequence 374 AA;
 Query Match 26.4%; Score 188; DB 14; Length 374;
 Best Local Similarity 71.2%; Pred. No. 3.5e-15;
 Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
 QY 73 RPLWIIIFSGNNNITKIKMPYIAGYKTFDGRRAEVSYYHNGAKFIRRVDCII 124
 DB 87 RPLWIIIFSGNNNITKIKMPYIAGYKTFDGRGAQYVYIGNGCPVFIRKVSNI 138

RESULT 10
 AAR45541
 ID AAR45541 standard; Protein; 374 AA.
 AC AAR45541;
 XX
 DT 13-JUL-1994 (first entry)
 XX
 DE Cry j I pollen allergen.
 XX
 KM Japanese cedar; detection; allergy; treatment; diagnosis;
 KM T cell epitope; sensitivity.
 XX
 OS Cryptomeria japonica.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..21
 FT /note= "signal peptide"
 FT 22..374
 FT Peptide /note= "mature peptide"
 XX
 PN WO9401560-A.
 XX
 PD 20-JAN-1994.
 XX
 PF 15-JAN-1993; 93WO-US00139.
 XX
 PR 01-SEP-1992; 92US-0938990.
 PR 10-JUL-1992; 92WO-US05661.
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 XX
 PI Bond JF, Garman RD, Griffith IU, Kuo M, Pollock J;
 XX
 DR WPI; 1994-035066/04.
 DR N-PSDB; AAQ55271.
 XX
 PT Antigens derived from Japanese cedar pollen allergen Cry j I -
 PT contain at least two T cell epitope(s), used to treat or diagnose
 PT allergy
 XX
 PS Disclosure, Fig 4; 137pp; English.

CC The sequence is that of the Japanese cedar pollen allergen
 CC Cry j 1 which contains at least two T cell epitopes. Peptide
 CC antigens derived from it can be used for the treatment and
 CC diagnosis of allergies associated with Japanese cedar pollen.
 CC The peptides have enhanced therapeutic properties but reduced
 CC side effects compared to naturally occurring allergens.

XX
 SQ Sequence 374 AA;

Query Match 26.4%; Score 188; DB 15; Length 374;
 Best Local Similarity 71.2%; Pred. No. 3.5e-15;
 Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

OY 73 RPLMTIFSGNNMNIKLMPTIAGYKTFDGRRAEVSVYHVNCAKIRRDGII 124
 DB 87 RPLMTIFSGNNMNIKLMPTIAGYKTFDGRGAQVYIGNGPCVFIRKVSNI 138

RESULT 11

AA82490
 ID AAR60166 standard; Protein; 374 AA.

XX
 AC AAR60166;

XX
 DT 24-MAR-1995 (first entry)

XX
 DE Japanese cedar pollen antigen Cryj1.

XX
 KW Japanese cedar; pollen antigen; allergen; Cryj1; sugi; pollinosis.

XX
 OS Crypomeria japonica.

XX
 FH Key location/Qualifiers

FT Peptide 1..21
 /label= signal_peptide

FT Protein 22..374
 /label= mature_Cryj1

XX
 JP06197768-A.

XX
 PD 19-JUL-1994.

XX
 PF 07-JAN-1993; 93JP-0001116.

XX
 PR 07-JAN-1993; 93JP-0001116.

XX
 PA (MEIJ) MEIJ SEIKA KAISHA.

XX
 DR WPI; 1994-268680/33.

XX
 DR N-PSDB; AAQ71601.

XX
 PT Sugi (Japanese cedar) pollen antigen Cryj1 - is useful for

XX
 PS diagnosis, treatment and prevention of sugi pollinosis

XX
 CC Claim 2; Page 5-7; 9pp; Japanese.

XX
 CC The coding sequence for the Japanese cedar ("sugi") pollen allergen

XX
 CC Cryj1 was isolated from a cDNA library prepared from polyA mRNA. All

XX
 CC or part of the Cryj1 protein can be used for diagnosis, treatment

XX
 CC and prevention of sugi pollinosis.

XX
 SQ Sequence 374 AA;

OY 73 RPLMTIFSGNNMNIKLMPTIAGYKTFDGRRAEVSVYHVNCAKIRRDGII 124

DB 87 RPLMTIFSGNNMNIKLMPTIAGYKTFDGRGAQVYIGNGPCVFIRKVSNI 138

RESULT 12

AA82490
 ID AAR82490 standard; Protein; 374 AA.

XX
 AC AAR82490;

XX
 DT 15-APR-1996 (first entry)

XX
 DE Cry j 1 Japanese Cedar pollen allergen.

XX
 KW Cry j 1; Japanese cedar pollen allergen; modified; drug production;

XX
 KW allergy; Crypomeria japonica.

XX
 OS Crypomeria japonica.

XX
 PN MO527786-A1.

XX
 PD 19-OCT-1995.

XX
 PF 06-APR-1995; 95MO-US04249.

XX
 PR 06-DEC-1994; 94US-0350225.

XX
 PR 08-APR-1994; 94US-0226248.

XX
 PA (TMCU-) IMMUNOLOGIC PHARM CORP.

XX
 PI Chen X, Evans S, Franzen HM, Kuo M, Powers SP;

XX
 PI Shaked Z;

XX
 DR WPI; 1995-366391/47.

XX
 DR N-PSDB; AAT04248.

XX
 PT Modified Crypomeria japonica (Cry j) 1 peptide(s) - useful for

XX
 PT treating allergy to Japanese cedar pollen allergen or

XX
 PT immunologically cross reactive allergens

XX
 PS Disclosure; Figure 1; 60pp; English.

XX
 CC Novel peptides of cry j 1 have been modified as a part of a

XX
 CC preformulation scheme to develop an optimised drug product for

XX
 CC therapeutic treatment of humans suffering from allergy to Japanese

XX
 CC cedar pollen allergen or an allergen which is immunologically cross

XX
 CC reactive with Japanese cedar pollen allergen. Such modified peptides

XX
 CC possess certain characteristics which render them particularly

XX
 CC suitable for drug product formulation. Peptide fragments of Cry j 1,

XX
 CC modified and unmodified, are given in AAR82491-R82525.

XX
 SQ Sequence 374 AA;

OY 73 RPLMTIFSGNNMNIKLMPTIAGYKTFDGRRAEVSVYHVNCAKIRRDGII 124

DB 87 RPLMTIFSGNNMNIKLMPTIAGYKTFDGRGAQVYIGNGPCVFIRKVSNI 138

RESULT 13

AA82565
 ID AAY25665 standard; Protein; 374 AA.

XX
 AC AAY25665;

XX
 DT 30-SEP-1999 (first entry)

XX
 DE Cedar allergen 493632 Cry j 1A protein fragment.

XX
 KW Major histocompatibility complex; class II; desensitising; human;

XX
 KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;

XX
 KW chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;

XX
 KW scree worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;

XX
 KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;

XX
 KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.

OS Cedrus sp.
 XX WO9934826-A1.
 PN 15-JUL-1999.
 PD
 XX 11-JAN-1999; 99WO-GB00080.
 PF 21-SEP-1998; 98GB-0020474.
 PR 09-JAN-1998; 98GB-0000445.
 XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 PA Kay AB, Larche M;
 PI WPI; 1999-458255/38.
 DR Desensitizing patients to polypeptide allergens
 XX Example 6, Page 73; 117pp; English.
 PS
 XX This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The methods can be used for
 CC desensitizing patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents a cedar (Cedrus sp.) allergen 493632 Cry j IA.
 XX
 SQ Sequence 374 AA;
 Query Match 26.4%; Score 188; DB 20; Length 374;
 Best Local Similarity 71.2%; Pred. No. 3.5e-15;
 Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
 QY 73 RPLWIFSGNNNIKLMKMPYIAGYKTFDGRRAEVSYYHVGAKFRRVDGI 124
 DB 87 RPLWIFSGNNNIKLMKMPYIAGYKTFDGRGAQYIYNGGPCVFIRKVSNT 138
 RESULT 14
 ID AAY25669 standard; protein; 374 AA.
 AC AAY25669;
 XX 30-SEP-1999 (first entry)
 DT Japanese cedar allergen 541802 Cry j I precursor protein fragment.
 XX Major histocompatibility complex; class II; desensitizing; human;
 KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; stinging;
 KW chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
 KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
 KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
 KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
 XX Cedrus sp.
 OS WO9934826-A1.
 PN 15-JUL-1999.

PF 11-JAN-1999; 99WO-GB00080.
 XX 21-SEP-1998; 98GB-0020474.
 PR 09-JAN-1998; 98GB-0000445.
 XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 PA Kay AB, Larche M;
 PI WPI; 1999-458255/38.
 DR Desensitizing patients to polypeptide allergens
 XX Example 6, Page 75; 117pp; English.
 PS
 XX This invention describes a novel method of desensitizing a patient to a
 CC polypeptide allergen and comprises administering to the patient a peptide
 CC derived from the allergen where restriction to a MHC Class II molecule
 CC possessed by the patient can be demonstrated for the peptide and the
 CC peptide is able to induce a late phase response in an individual who
 CC possesses the MHC Class II molecule. The methods can be used for
 CC desensitizing patients to allergens present in e.g. grass, tree and weed
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
 CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
 CC produce immunological vaccines which may be used to prevent and/or treat
 CC conditions involving hypersensitivity to allergens. This sequence
 CC represents a Japanese cedar (Cedrus sp.) allergen 541802 Cry j I
 XX
 SQ Sequence 374 AA;
 Query Match 26.4%; Score 188; DB 20; Length 374;
 Best Local Similarity 71.2%; Pred. No. 3.5e-15;
 Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
 QY 73 RPLWIFSGNNNIKLMKMPYIAGYKTFDGRRAEVSYYHVGAKFRRVDGI 124
 DB 87 RPLWIFSGNNNIKLMKMPYIAGYKTFDGRGAQYIYNGGPCVFIRKVSNT 138
 RESULT 15
 ID AAB69103 standard; Protein; 210 AA.
 AC AAB69103;
 XX 23-APR-2001 (first entry)
 DT Cedar pollen allergen T cell epitope derived protein SBQ ID NO:12.
 XX Japanese cedar; Cryptomeria japonica; cedar pollen allergen;
 KW T cell epitope; anti-stinging pollinosis.
 XX Cryptomeria japonica.
 OS Synthetic.
 XX JP2000327699-A.
 PN 28-NOV-2000.
 PD 15-MAR-2000; 2000JP-0071710.
 PF 15-MAR-1999; 99JP-0068316.
 PR (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA (SANY) SANKYO CO LTD.
 XX WPI; 2001-185061/19.
 DR N-PSDB; AAF59012.

XX Novel peptide and its use -

PS Claim 7; Page 39-40; 75pp; Japanese.

CC The present invention describes a peptide, its complex, derivative or
CC its polymerizate, where the peptide (I) has a formula of:
CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7
CC where alpha1 to alpha7 = amino acid sequence selected from amino acid
CC sequences ranging from 11-19 amino acids derived from T cell epitopes
CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen
CC allergens. The peptide can be used in an antiasthma agent.
CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used
CC in the exemplification of the present invention.

SO Sequence 210 AA;

Query Match 25.0%; Score 178; DB 22; Length 210;

Best Local Similarity 32.5%; Pred. No. 3e-14; Mismatches 7; Indels 96; Gaps 6;

Db 1 MKVTVAENQPGPRRVFIRGVNVIHGRIDIFASKNFHLQNT-IGTGR----- 50
29 MKVTVAENQPGP-----FASKNFHLQNTYKLTSGKIASCLNTYGL 67
51 -----RWKNNRIWL-----QFALTGFTLMGRRPLMIIFSGNMNIIKLKMP 90
68 VHVANNNDPSGKYEGNGIYTKKEAFNVEQFALTGFTLMGRR----- 110
QY 91 MYTAGYKTFDGRPAEVSIVHNGAKFIRRVDTGIIAAYONPASM 133
111 -----DPR-----GIIAAYONPASM 125

Search completed: April 23, 2003, 18:07:49
Job time : 38 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:25 ; Search time 15.5789 Seconds

(without alignments)
212.987 Million cell updates/sec

Title: US-09-142-524D-1

Perfect score: 406

Sequence: 1 MKYTVAFNPGPNRRVFIRK.....IASRVYDGIATAYQNPASWK 80

Scoring table: BLOSUM62

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99	24.4	514	1	MPA2_CRYJA
2	74	18.2	375	1	MPA1_CHAOB
3	73	18.0	346	1	MPA1_CUPAR
4	73	18.0	367	1	MPA1_JUNAS
5	72	17.7	234	1	A29B_DROME
6	71	17.5	374	1	SBP_CRYJA
7	70	17.2	152	1	RS18_HUMAN
8	66.5	16.4	196	1	YMO7_YEAST
9	62.5	15.4	145	1	YPI2_AGRU
10	62.5	15.4	434	1	PEL_TILLO
11	62.5	15.4	1290	1	PIG1_HUMAN
12	62	15.3	173	1	PAPB_ECOLI
13	61.5	15.1	633	1	TOPI_THEMA
14	61	15.0	152	1	RS18_DROME
15	60.5	14.9	110	1	Y12K_SMSV4
16	60	14.8	727	1	PKP1_BOVIN
17	60	14.8	747	1	PKP1_HUMAN
18	59.5	14.7	842	1	PHS2_RAT
19	59	14.5	152	1	RS18_SCHPO
20	59	14.5	234	1	A29B_DROME
21	59	14.5	840	1	MURS_LACTA
22	59	14.5	862	1	PKM2_YEAST
23	58.5	14.4	478	1	YSP3_YEAST
24	58.5	14.4	842	1	PHS2_YEAST
25	58.5	14.4	1291	1	PIG1_BOVIN
26	58	14.3	152	1	RS18_ARATH
27	58	14.3	565	1	HEMA_IADAI
28	58	14.3	646	1	NODO_RHIS3
29	58	14.3	728	1	PKP1_MOUSE
30	58	14.3	1070	1	YH4_YEAST
31	57.5	14.2	207	1	PIG1_MOUSE
32	57.5	14.2	1290	1	PIG1_RAT
33	57	14.0	379	1	YLEU_DEBOC

34	57	14.0	511	1	UDPGL_DICDI
35	57	14.0	567	1	ODP2_HAEIN
36	56.5	13.9	99	1	IHEA_XYLEA
37	56.5	13.9	215	1	RAN_BRUMA
38	56.5	13.9	339	1	BMDA_BORBU
39	56.5	13.9	394	1	MEMB_METTR
40	56.5	13.9	842	1	PHS2_HUMAN
41	56.5	13.9	994	1	PPOL_DROME
42	56.5	13.9	1163	1	PROD_PEA
43	56	13.8	246	1	RS3_AERPE
44	56	13.8	267	1	RS3_MYCCE
45	56	13.8	341	1	AAR_AQUUP

ALIGNMENTS

RESULT 1

MPA2_CRYJA

ID MPA2_CRYJA STANDARD: PRT, 514 AA.

AC P43212;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Possible polygalacturonase precursor (BC 3.2.1.15) (PG) (Pectinase)

OS (Major pollen allergen Cry j 2) (Cry j II).

OS Cryptomeria japonica (Japanese cedar).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.

OX NCBI_TaxID=3369;

[1]

SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.

RC TISSUE=Pollen;

RA MEDLINE=95010777; PubMed=7926035;

RA Namba M., Kurose M., Torigoe K., Hino K., Taniguchi Y., Fukuda S.,

RA Usui M., Kurimoto M.;

RT "Molecular cloning of the second major allergen, Cry j II, from

RT Japanese cedar pollen.";

RT FEBS Lett. 353:124-128(1994).

[2]

SEQUENCE FROM N.A.

RC TISSUE=Pollen;

RA MEDLINE=94271186; PubMed=8002972;

RA Komiyama N., Sone T., Shimizu K., Morikubo K., Kino K.;

RT "cDNA cloning and expression of Cry j II the second major allergen of

RT Japanese cedar pollen.";

RT Biochem. Biophys. Res. Commun. 201:1021-1028(1994).

[3]

SEQUENCE OF 55-64.

RA MEDLINE=90342988; PubMed=2382797;

RA Satauchi M., Inouye S., Tanai M., Ando S., Usui M., Matuhasi T.;

RT "Identification of the second major allergen of Japanese cedar

RT pollen.";

RT Allergy 45:309-312(1990).

-1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-

-1- galactosiduronic linkages in pectate and other galacturonans.

-1- SUBCELLULAR LOCATION: SECRETED OR AMYLOPLAST (POTENTIAL).

-1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES

(POLYGALACTURONASES).

DR PROSITE, PS00502; POLYGALACTURONASE; 1.
 KM HYDROLASE; Glycosidase; Cell wall; Signal; Zymogen; Fruit ripening;
 FT AMYLOPLAS; Glycoprotein; Allergen.
 FT SIGNAL 1 22
 FT PROPEP 23 45
 FT CHAIN 46 433
 FT PROPEP 434 514
 FT ACT SITE 278 278
 FT CARBOHYD 460 460
 FT CARBOHYD 472 472
 FT CONFLICT 5 5
 FT CONFLICT 12 12
 FT CONFLICT 34 35
 FT CONFLICT 37 37
 FT CONFLICT 88 88
 FT CONFLICT 98 98
 FT CONFLICT 451 451
 FT CONFLICT 454 454
 FT CONFLICT 504 504
 FT CONFLICT 507 507
 SQ SEQUENCE 514 AA; 56645 MW; 624611C3FA8D6302 CRC64;
 Query Match 24.4%; Score 99; DB 1; Length 514;
 Best Local Similarity 100.0%; Pred. No. 0.00031;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 IDIPASKNFHLOKNTIGTG 49
 DB 236 IDIPASKNFHLOKNTIGTG 254

RESULT 2
 MPAL CHAOS STANDARD; PRT; 375 AA.
 AC 096385;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Major pollen allergen (Japanese cypress).
 OS Chamaecyparis obtusa (Japanese cypress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae;
 OC Chamaecyparis.
 NCBI_TaxID=13415;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Pollen;
 RX MEDLINE=96255194; PubMed=8676896;
 RA Suzuki M., Komiyama N., Itoh M., Sone T., Kuno K., Takagi I.,
 RT "Purification, characterization and molecular cloning of Cha o 1, a
 RT major allergen of Chamaecyparis obtusa (Japanese cypress) pollen."
 RL Mol. Immunol. 33:451-460(1996).
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; DA5404; BAA08246.1; -
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 DR PRINTS; PRO0807; AMBALLERGEN.
 KW Allergen; Glycoprotein; Signal.
 FT CHAIN 1 21
 FT SIGNAL 22 375
 FT CARBOHYD 110 110
 FT CARBOHYD 148 148
 MAJOR POLLEN ALLERGEN CHA O 1.
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 178 178
 FT CARBOHYD 293 293
 FT CARBOHYD 352 352
 SQ SEQUENCE 375 AA; 40258 MW; 81CD91DF06EDBBF CRC64;
 Query Match 18.2%; Score 74; DB 1; Length 375;
 Best Local Similarity 59.3%; Pred. No. 0.21;
 Matches 16; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

OY 1 MKVTYAFNFGPNRVRFIKRVSVIHH 27
 DB 234 MKVTYAFNFGPNRVRFIKRVSVIHH 260

RESULT 3
 MPAL CUPAR STANDARD; PRT; 346 AA.
 AC 096389;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Major pollen allergen Cup a 1.
 OS Cupressus arizonica.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
 NCBI_TaxID=49011;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20571526; PubMed=11122214;
 RA Acetluno E., Del Pozo V., Minguez A., Arrieta I., Cortegano I.,
 RA Cardaba B., Gallardo S., Rojo M., Palomino P., Lahoz C.;
 RT "Molecular cloning of major allergen from Cupressus arizonica pollen:
 RT Cup a 1."
 RL Clin. Exp. Allergy 30:1750-1758 (2000).
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AJ243570; CAB62551.1; -
 DR InterPro; IPR002022; Amb allergen.
 DR Pfam; PF00544; pec_lyase; 1.
 DR PRINTS; PRO0807; AMBALLERGEN.
 KW Allergen; Glycoprotein.
 FT CARBOHYD 127 127
 FT CARBOHYD 157 157
 FT CARBOHYD 272 272
 SQ SEQUENCE 346 AA; 37589 MW; F1281DDDA1D5DDF0 CRC64;
 Query Match 18.0%; Score 73; DB 1; Length 346;
 Best Local Similarity 55.6%; Pred. No. 0.25;
 Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

OY 1 MKVTYAFNFGPNRVRFIKRVSVIHH 27
 DB 213 MKVTYAFNFGPNRVRFIKRVSVIHH 239

RESULT 4
 MPAL JUNAS STANDARD; PRT; 367 AA.
 AC P81294; Q9ZNT7;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Major pollen allergen Jun a 1 precursor.
 OS Juniperus ashei (Ozark white cedar).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
 OX NCB1_TaxID=13101;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 31-38; 42-50; 58-80; 88-94;
 RC 117-124; 134-140; 160-164; 256-263 AND 322-325.
 RA TISSUE=Pollen;
 RX MEDLINE=99414163; PubMed=10482836;
 RA Miodoro-Horvut T.M., Goldblum R.M., Kurosky A., Wood T.G.,
 RA Brooks E.G.;
 RT "Molecular cloning of mountain cedar (Juniperus ashei) pollen major
 RT allergen, Jun a 1.",
 RL J. Allergy Clin. Immunol. 104:613-617(1999).
 RN [2]
 RP SEQUENCE OF 22-50.
 RC TISSUE=Pollen;
 RX MEDLINE=99414162; PubMed=10482835;
 RA Miodoro-Horvut T., Goldblum R.M., Kurosky A., Goetz D.W.,
 RA Brooks E.G.;
 RT "Isolation and characterization of the mountain cedar (Juniperus
 RT ashei) pollen major allergen, Jun a 1.",
 RL J. Allergy Clin. Immunol. 104:608-612(1999).
 CC -1- DISEASE: THIS PROTEIN CAUSES SEVERAL SEASONAL ALLERGIC RHINITIS
 CC IN NORTH AMERICA.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC -1- AMB A I/AMB A II/CRY J I SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF106663; AAD03609.1; -;
 CC EMBL; AF106662; AAD03608.1; -;
 CC InterPro: IPR002022; Amb_allergen.
 CC Pfam: PF00544; pec_lyase; 1.
 CC PRINTS; PR00807; AMBALLERGEN.
 CC Allergen; Glycoprotein; Signal.
 CC SIGNALL 21
 CC FT CHAIN 1 21 MAJOR POLLEN ALLERGEN JUN A 1.
 CC FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SQ SEQUENCE 367 AA; 39824 MW; FC9B81E675662E49 CRC64;
 CC -----
 CC Query Match 18.0%; Score 73; DB 1; Length 367;
 CC Best Local Similarity 55.6%; Pred. No. 0.27;
 CC Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
 CC -----
 CC QY 1 MKTVAFNPGGPNRRVFIKRVNVIIH 27
 CC DB 234 MKTVAFNPGGPNRRVFIKRVNVIIH 260
 CC -----
 CC RESULT 5
 CC A29B DROME STANDARD; PRT; 234 AA.
 CC AC 046197; Q9TV73; Q9TW05; Q9TW07; Q9U976; Q9U977; Q9U978;
 CC AC Q9U979; Q9V305;
 CC DT 16-OCT-2001 (Rel. 40, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Accessory gland protein Acp29AB precursor.
 CC GN ACP29AB OR CGI1797.
 CC OS Drosophila melanogaster (Fruit fly).
 CC OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 CC OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 CC OX NCB1_TaxID=7227;
 CC RN [1]
 CC RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.

RC STRAIN=Canton-S; TISSUE=Male accessory gland;
 RX MEDLINE=98135120; PubMed=9474779;
 RA Wolfner M.F., Harada H.A., Berman M.J., Stelick T.J., Kraus K.W.,
 RA Kalb J.M., Lung Y.O., Neubaum D.M., Park M., Tram U.K.;
 RT "New genes for male accessory gland proteins in Drosophila
 RT melanogaster.",
 RL Insect Biochem. Mol. Biol. 27:825-834(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Various strains;
 RX MEDLINE=99282496; PubMed=10353898;
 RA Aguade M.;
 RT "Positive selection drives the evolution of the Acp29AB accessory
 RT gland protein in Drosophila.",
 RL Genetics 152:543-551(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Various strains;
 RX MEDLINE=20556153; PubMed=11102381;
 RA Begun D.J., Whitely P., Todd B.L., Waldrip-Dall H.M., Clark A.G.;
 RT "Molecular population genetics of male accessory gland proteins in
 RT Drosophila.",
 RL Genetics 156:1879-1888(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amaratunga C., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer V., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abul J.F., Agbayani A., An H.-D., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,
 RA Butte K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mochlyov G., Molehina N.V., Mobarry C., Morris J., Moshireff A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J.D., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.E., Smith T.,
 RA Stryker R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasatman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Zhu H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.",
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: RESPONSIBLE FOR PHYSIOLOGICAL AND BEHAVIORAL CHANGES IN
 CC MATED FEMALE FLIES.
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -1- TISSUE SPECIFICITY: MAIN CELLS OF THE ACCESSORY GLAND AND IN
 CC SEMINAL FLUID.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

DR	EMBL	U85758; AAB95382.1; -
DR	EMBL	AJ240513; CAB53187.1; -
DR	EMBL	AJ240514; CAB53188.1; -
DR	EMBL	AJ240515; CAB53189.1; -
DR	EMBL	AJ240516; CAB53190.1; -
DR	EMBL	AJ240517; CAB53191.1; -
DR	EMBL	AJ240518; CAB53192.1; -
DR	EMBL	AJ240519; CAB53193.1; -
DR	EMBL	AJ240520; CAB53194.1; -
DR	EMBL	AJ240521; CAB53195.1; -
DR	EMBL	AJ240522; CAB53196.1; -
DR	EMBL	AJ240523; CAB53197.1; -
DR	EMBL	AJ240524; CAB53198.1; -
DR	EMBL	AJ240525; CAB53199.1; -
DR	EMBL	AJ240526; CAB53200.1; -
DR	EMBL	AJ240527; CAB53201.1; -
DR	EMBL	AJ240528; CAB53202.1; -
DR	EMBL	AJ240529; CAB53203.1; -
DR	EMBL	AJ240530; CAB53204.1; -
DR	EMBL	AJ240531; CAB53205.1; -
DR	EMBL	AJ240532; CAB53206.1; -
DR	EMBL	AJ240533; CAB53207.1; -
DR	EMBL	AJ240534; CAB53208.1; -
DR	EMBL	AJ240535; CAB53209.1; -
DR	EMBL	AJ240536; CAB53210.1; -
DR	EMBL	AJ240537; CAB53211.1; -
DR	EMBL	AJ240538; CAB53212.1; -
DR	EMBL	AJ240539; CAB53213.1; -
DR	EMBL	AJ240540; CAB53214.1; -
DR	EMBL	AJ240541; CAB53215.1; -
DR	EMBL	AJ240542; CAB53216.1; -
DR	EMBL	AJ240543; CAB53217.1; -
DR	EMBL	AJ240544; CAB53218.1; -
DR	EMBL	AJ240545; CAB53219.1; -
DR	EMBL	AJ240546; CAB53220.1; -
DR	EMBL	AJ240547; CAB53221.1; -
DR	EMBL	AJ240548; CAB53222.1; -
DR	EMBL	AJ240549; CAB53223.1; -
DR	EMBL	AJ240550; CAB53224.1; -
DR	EMBL	AJ240551; CAB53225.1; -
DR	EMBL	AY010527; AAG32559.1; -
DR	EMBL	AY010528; AAG32560.1; -
DR	EMBL	AY010529; AAG32561.1; -
DR	EMBL	AY010530; AAG32562.1; -
DR	EMBL	AY010531; AAG32563.1; -
DR	EMBL	AY010532; AAG32564.1; -
DR	EMBL	AY010533; AAG32565.1; -
DR	EMBL	AY010534; AAG32566.1; -
DR	EMBL	AY010535; AAG32567.1; -
DR	EMBL	AY010536; AAG32568.1; -
DR	EMBL	AY010537; AAG32569.1; -
DR	EMBL	AY010538; AAG32570.1; -
DR	EMBL	AY010539; AAG32571.1; -
DR	EMBL	AY010540; AAG32572.1; -
DR	EMBL	AY010541; AAG32573.1; -
DR	EMBL	AY010542; AAG32574.1; -
DR	EMBL	AY010543; AAG32575.1; -
DR	EMBL	AE003621; AAF52665.1; -
DR	FLYBase	FB90015583; ACP29AB.
DR	Interpro	IPR001304; Lectin_C.
DR	Pfam	PF00059; Lectin_c; 1.
DR	PROSITE	PS00615; C_TYPE_LLECTIN_1; FALSE_NEG.
DR	PROSITE	PS00641; C_TYPE_LLECTIN_2; 1.
KM	Behavior	Lectin; Signal; Polymorphism.
FT	Signal	1
CHAIN	1	234
	22	234
	21	POTENTIAL.
		ACCESSORY GLAND PROTEIN ACP29AB.

FT	D	DOMAIN	137	228	C-TYPE LECTIN (LONG FORM).
FT	C	CARBOHYD	61	61	N-LINKED (GLCNAC...) (POTENTIAL).
FT	C	CARBOHYD	164	164	N-LINKED (GLCNAC...) (POTENTIAL).
FT	V	VARIANT	29	29	K -> N (IN STRAINS MA7, MS16 AND MS26).
FT	V	VARIANT	59	59	K -> R (IN STRAINS LA1, LA3, LA4, LA5,
FT	L	LA14, LA15, LA16, LA21, LA25, LA34, LA35,			
FT	M	MO1B, MO52B, MO80B, MA5, MA8, MA13, MA18,			
FT	M	MA20, MA21, MA45, MA48, MA52, MA67,			
FT	Z	ZIM26, ZIM29, ZIM30, ZIM56, MS6 AND			
FT	M	MS19).			
FT	A	-> S (IN STRAINS LA16 AND LA35).			
FT	R	-> L (IN STRAINS LA14, LA14, LA16,			
FT	L	LA35, MO2B, MO15B, MO34A, MO36A, MO37A,			
FT	M	MO47A, MO52B, MO58B, MO80B, MA5, MA7,			
FT	M	MA13, MA18, MA20, MA21, MA45, MA48, MA50,			
FT	M	MA52, MA67, MS1, MS6, WS9, WS12, MS16,			
FT	W	WS26, WS47, MS56, ZIM2, ZIM25, ZIM29,			
FT	Z	ZIM30, ZIM32, ZIM37, ZIM42 AND ZIM56).			
FT	K	-> M (IN STRAINS BERKELEY, LA13, LA14,			
FT	L	LA15, LA16, MO1B, MO8B, MO40B, MO52B,			
FT	M	MA5, MA21, MA45, MA52, MA67, ZIM29,			
FT	Z	ZIM30, ZIM42 AND ZIM56).			
FT	E	-> D (IN STRAIN ZIM57).			
FT	S	SEQUENCE	234 AA;	214 MM;	9854QCDJ215460F3E CRC64;
FT	V	VARIANT	214	214	
FT	S	SEQUENCE	234 AA;	27173 MM;	9854QCDJ215460F3E CRC64;
Qy		Query Match	17.7%;	Score 72;	DB 1; Length 234;
Db		Best Local Similarity	25.9%;	Pred. No.	0.22;
Oy		Matches	21;	Conservative	19;
Db				Mismatches	31;
Oy				Indels	10;
Db				Gaps	3
Oy	9	QFGNRRVFLIKRSNV----	IIGHGRIDIFASKNHLOKNTGT--GRRISLKTSGSKIA	62	
Db	88	QLQPLKLTIMHHASNIRKSNITKRREFEKVSRRPHIKNLMTWFPAVYTCRGNCHLA	147		
Oy	63	S-----RRVDGIILAYONPASW	79		
Db	148	NIDPEKELDGIALAPNNSTW	168		
RESULT 6					
ID	SBP-CRYUA	STANDARD;	PRT;	374 AA.	
AC	P18632;				
DT	01-NOV-1990	(Rel. 16, Created)			
DT	01-NOV-1995	(Rel. 32, Last sequence update)			
DT	30-MAY-2000	(Rel. 39, Last annotation update)			
DS	Sbp basic protein precursor (SBP) (Major allergen Cry j I) (Cry j I).				
OC	Cryptomeria japonica (Japanese cedar); Embryophyta; Tracheophyta;				
OC	Eukaryota; Viridiplantae; Streptophyta; Confariales; Cupressaceae; Cryptomeria.				
NCBI	_TaxID=3369;				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RC	TISSUE=Pollen;				
RC	MEDLINE=94I83234; PubMed=8135802;				
RA	Sone T., Komiyama N., Shimizu K., Kusakabe T., Morikubo K.,				
RA	Kino K.;				
RT	"Cloning and sequencing of cDNA coding for Cry j I, a major allergen				
RL	of Japanese cedar pollen.";				
RN	Biochem. Biophys. Res. Commun. 199;619-625(1994).				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Pollen;				
RC	Namba M., Kurose M., Torigoe K., Fukuda S., Kurimoto M.;				
RA	Tanaihi M., Ando S., Ueui M., Kurimoto M., Sakaguchi M., Inoue S.,				
RA	Matsumae T.;				
RT	"N-terminal amino acid sequence of a major allergen of Japanese cedar				
RL	pollen (Cry j I).";				
RN	Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.				
RP	SEQUENCE OF 22-41.				
RC	TISSUE=Pollen;				
RC	MEDLINE=89031257; PubMed=3181435;				
RA	Tanaihi M., Ando S., Ueui M., Kurimoto M., Sakaguchi M., Inoue S.,				
RA	Matsumae T.;				
RT	"N-terminal amino acid sequence of a major allergen of Japanese cedar				
RL	pollen (Cry j I).";				
RN	FEB8 Lett. 239;329-332(1988).				

[4]
 RN CARBOHYDRATES.
 RP TISSUE-Pollen.
 RC MEDLINE=95003748; PubMed=7920021;
 RA Hijikata A., Matsumoto I., Kojima K., Ogawa H.;
 RT "Antigenicity of the oligosaccharide moiety of the Japanese cedar
 RL Int. Arch. Allergy Immunol. 105:198-202(1994).
 RN [5]
 RP STRUCTURE OF CARBOHYDRATES.
 RC TISSUE-Pollen;
 RX MEDLINE=95332249; PubMed=7608114;
 RA Hino K., Yamamoto S., Sano O., Taniguchi Y., Kohno K., Usui M.,
 RT Fukuda S., Hanzawa H., Hanyama H., Kurimoto M.;
 RT Carbohydrate structures of the glycoprotein allergen Cry j 1 from
 RL Japanese cedar (Cryptomeria japonica) pollen.";
 CC Biochem. 117:289-295(1995)
 CC -1- PPM: CONTAINS PUCCOSE/XULOSE-CONTAINING N-LINKED OLIGOSACCHARIDES.
 CC -1- DISEASE: THIS PROTEIN IS THE MAJOR ALLERGEN OF JAPANESE CEDAR
 CC POLLEN, THE MOST COMMON POLLEN ALLERGEN IN JAPAN.
 CC -1- MISCELLANEOUS: THE SEQUENCE OF CRY J 1 FORM A IS SHOWN HERE. FORM
 CC B DIFFERS IN SIX POSITIONS.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D26544; BAA05542.1; -
 DR EMBL; D26545; BAA05543.1; -
 DR EMBL; D34639; BAA07020.1; -
 DR PIR; A44773; A44773.
 DR GlycoSuiteDB: P18632; -
 DR InterPro: IPR002022; Amb_allergen.
 DR Pfam: PF00544; pec_lyase; 1.
 DR PRINTS; PR00807; AMBALLERGEN.
 KW Allergen; Glycoprotein; Multigene family; Signal.
 FT SIGNAL 1 21
 FT CHAIN 1 374
 FT CARBOHYD 158 158
 FT CARBOHYD 191 191
 FT CARBOHYD 293 293
 FT CARBOHYD 354 354
 FT VARIANT 12 12
 FT VARIANT 143 143
 FT VARIANT 202 202
 FT VARIANT 221 221
 FT VARIANT 358 358
 FT VARIANT 361 361
 SQ SEQUENCE 374 AA, 40645 MW, 74AB2595024856F CRC64;
 Query Match 17.5%; Score 71; DB 1; Length 374;
 Best Local Similarity 55.6%; Pred. No. 0.48;
 Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

1 MKVTAFNFGNRRVFIKRVSNVITH 27
 |||||
 DB 234 MKVTAFNFGNRRVFIKRVSNVITH 260

RESULT 7
 RS18 HUMAN STANDARD; PRT; 152 AA.
 AC P2532; -
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 15-UN-2002 (Rel. 41, Last annotation update)
 DE 40S ribosomal protein S18 (KE-3) (KE3).
 GN RPS18.
 OS Homo sapiens (Human).
 OS Mus musculus (Mouse).
 OS Rattus norvegicus (Rat), and
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI Taxid=9606, 10090, 10116, 9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Human; TISSUE=Placenta;
 RX MEDLINE=93181276; PubMed=8441687;
 RA Chassin D., Bellet D., Roman A.;
 RT "The human homolog of ribosomal protein S18.";
 RL Nucleic Acids Res. 21:745-745(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Human;
 RA Tubby B.;
 RT Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Rat; STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=91337062; PubMed=1872840;
 RA Chan Y.-L., Paz V., Wool I.G.;
 RT "The primary structure of rat ribosomal protein S18.";
 RL Biochem. Biophys. Res. Commun. 178:1212-1218(1991).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Mouse;
 RX MEDLINE=92182530; PubMed=1543907;
 RA Macmurray A.J., Shin H.S.;
 RT "The murine MHC encodes a mammalian homolog of bacterial ribosomal
 RL protein S13.";
 RN Mamm. Genome 2:87-95(1992).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Mouse; STRAIN=129/SvJ;
 RA Rowen L., Qin S., Madan A., Loretz C., James R., Dors M., Mix L.,
 RA Hall J., Lasky S., Hood L.;
 RT "Sequence of the mouse major histocompatibility locus class II
 RL region.";
 RN Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Pig;
 RA Kimura M., Kawakami K., Suzuki H., Hamajima N.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN THE BINDING OF FMET-TRNA AND, HENCE, IN THE
 CC INITIATION OF TRANSLATION (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE S13P FAMILY OF RIBOSOMAL PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X57529; CAA0750.1; -
 DR EMBL; AL031228; CAA20231.1; -
 DR EMBL; X69150; CAB56794.1; -
 DR EMBL; W76763; AAA16795.1; -
 DR EMBL; W76762; AAA16796.1; -
 DR EMBL; AF100956; AAC69898.1; -
 DR EMBL; AF110520; AAC97978.1; -
 DR EMBL; AB000911; BAA19211.1; -
 DR PIR; JH0419; R3RT18.
 DR PIR; S30393; S30393.
 DR Genew; HGNC:10401; RPS18.

MM; 180473; -.
 DR MGD; MGI:98146; Rps18.
 DR InterPro; IPR001892; Ribosomal_S13.
 DR Pfam; PF00416; Ribosomal_S13; 1.
 DR ProDom; PD001363; Ribosomal_S13; 1.
 DR PROSITE; PS00646; RIBOSOMAL_S13; 1.
 KW Ribosomal protein.
 FT CONFLICT 104 104 D -> H (IN REF. 4; AAL6796).
 SQ SEQUENCE 152 AA; 17719 MW; 4ADAF662C3F37F22 CRC64;
 Query Match 17.2%; Score 70; DB 1; Length 152;
 Best Local Similarity 30.4%; Pred. No. 0.23;
 Matches 24; Conservative 7; Mismatches 32; Indels 16; Gaps 2;

OY 12 PNRRPFRKVSNTIHERRIDIFASKPHLOKNTIGTGR-----ISLKTSGKI 61
 DB 6 PEKFGHILRVNTNIDGRRKIAFALTAK-----GVGRYAHVVLKRAIDLTFRAGEL 59
 OY 62 ASRAVDGIIAAYQNPASWK 80
 DB 60 TEDEVRYITIMQNPQYK 78

RESULT 8
 YMO7 YEAST STANDARD; PRT; 196 AA.
 AC 004487;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Putative succinate dehydrogenase cytochrome B subunit, mitochondrial precursor.
 GN YMR18C OR YMR718.17C.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxId=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288C / AB972;
 RA Hunt S., Bowman S., Barrell B.G., Rajandream M.A.;
 RL Submitted (May-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MONO-HEME CYTOCHROME B. INVOLVED IN SYSTEM II OF THE
 CC MITOCHONDRIAL ELECTRON TRANSPORT CHAIN WHICH IS RESPONSIBLE FOR
 CC TRANSFERRING ELECTRONS FROM SUCCINATE TO UBIQUINONE (COENZYME Q)
 CC (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B560 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z49702; CAAG9756.1; -.
 CC SGD; S0004724; YMR18C.
 DR InterPro; IPR000701; Sdh_cyt.
 DR Pfam; PF01127; Sdh_cyt; 1.
 DR PROSITE; PS01000; SDH_CYT_1; 1.
 DR PROSITE; PS01001; SDH_CYT_2; 1.
 KW Hypothetical protein; Tricarboxylic acid cycle; Electron transport;
 KW Heme; Transmembrane; Mitochondrion; Transmembrane;
 FT TRANSIT 1 2 MITOCHONDRION (POTENTIAL).
 FT CHAIN 1 196 PUTATIVE SUCCINATE DEHYDROGENASE
 FT CYTOCHROME B SUBUNIT.
 FT TRANSMEM 99 119 POTENTIAL.
 FT TRANSMEM 175 195 POTENTIAL.
 RC STRAIN=cv. Neillie white; TISSUE=Pollen;
 SQ SEQUENCE 196 AA; 22309 MW; 4143398B9B2B057 CRC64;

Query Match 16.4%; Score 66.5; DB 1; Length 196;
 Best Local Similarity 29.7%; Pred. No. 0.81;
 Matches 27; Conservative 16; Mismatches 33; Indels 15; Gaps 7;

OY 1 MKVNV--AFNPG--PNRRVFIKVS-----NVIIHGRRIDIFASKPH--LOKNTIGT 48
 DB 1 MKATIQRTSVFVGPRASVFPRISTPILHNTYISNG-RMDLF-SKEFHNQSVSKSLMS 58

OY 49 GRISLKTSGKIASRRVDGIIAAYQNPASW 79
 DB 59 SNKEELLVSQR-KRPIPSPHLVTEPEMSW 88

RESULT 9
 YP12 AGRTU STANDARD; PRT; 145 AA.
 AC 044433;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 15.8 kDa protein in pinF2.3' region (ORF2).
 OS Agrobacterium tumefaciens.
 OG Plasmid pTiA6.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxId=358;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89213933; PubMed=2708311;
 RA Kanemoto R.H., Powell A.T., Akiyoshi D.B., Regier D.A.,
 RA Kerstetter R.A., Neeter E.W., Hawes M.C., Gordon M.P.;
 RT Nucleotide sequence and analysis of the plant-inducible locus pinF
 RT from Agrobacterium tumefaciens.
 RL J. Bacteriol. 171:2506-2512 (1989).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M19352; AA82504.1; -.
 KW Hypothetical protein; Plasmid.
 KM
 SQ SEQUENCE 145 AA; 15829 MW; 95A3D3D8B560A6 CRC64;

Query Match 15.4%; Score 62.5; DB 1; Length 145;
 Best Local Similarity 33.3%; Pred. No. 1.7;
 Matches 17; Conservative 11; Mismatches 18; Indels 5; Gaps 2;

OY 18 IKRVSNVTH-GRRIDIFASKPHLOKNTIGTGRKRSLSLKTSGKIASRRVD 67
 DB 1 MKRISTIVGVFLATPYVADNTH---TLGTLSEIELALTAKPVNVTV 47

RESULT 10
 PEL_LILLO STANDARD; PRT; 434 AA.
 AC P40973;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Pectate lyase precursor (BC 4.2.2.2).
 OS Lilium longiflorum (Trumpet Lily).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae;
 CC Lilium.
 OX NCBI_TaxId=4690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Neillie white; TISSUE=Pollen;
 RA Kim S.R., Finkel D.J., An G.;

RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Eliminative cleavage of pectate to give
 CC oligosaccharides with 4-deoxy-alpha-D-gluc-4-enurosyl groups at
 CC their non-reducing ends.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z17328; CAAT8976.1; -;
 DR EMBL; L18911; AAA33398.1; -;
 DR PIR; S29612; S29612.
 DR InterPro: IPR002023; Amb allergen.
 DR Pfam: PF00544; Pect_lyase_1.
 DR PRINTS; PR00807; AMBALLERGEN.
 KM Lyase; Signal.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 434 PECTATE LYASE.
 FT ACT SITE 312 312 POTENTIAL.
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 434 AA; 48457 MW; C1F3E30AD2BBD064 CRC64;
 Query Match 15.4%; Score 62.5; DB 1; Length 434;
 Best Local Similarity 28.4%; Pred. No. 5.8; Indels 19; Gaps 4;
 Matches 25; Conservative 11; Mismatches 33;
 QY 1 MKYTVANFGNGEN-----RRVFKRVSNVTHGRRRIDIFASKNPHLQKNTIGTGR- 51
 DB 296 MQYTVANFNHGRGLVGRMPRCRGYFVHVNDYTHIMVAVGSGR-----PTISGKNY 351
 QY 52 ISLKITSGKIASRRVDGIIAAYGNPASW 79
 DB 352 IAPHIEAKEVTGR-----DYAEPAEW 373
 RESULT 11
 P1GI1_HUMAN STANDARD; PRT; 1290 AA.
 AC P19174;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase gamma 1
 DE (EC 3.1.4.11) (PLC-gamma-1) (Phospholipase C-gamma-1) (PLC-II)
 GN PLCG1 OR PLC1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eumetazoa; Primates; Carnivora; Hominoidea; Homo.
 OX NCBI_TaxID:9606;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Vein, and Brain;
 RX MEDLINE=9035593; PubMed=2167438;
 RA Burgess W.H., Dione C.A., Kaplow J.M., Mudd R., Friesel R.,
 RA Zilberstein A., Schlesinger J., Jaye M.;
 RT "Characterization and cDNA cloning of phospholipase C-gamma, a major
 RT substrate for heparin-binding growth factor I (acidic fibroblast
 RT growth factor)-activated tyrosine kinase.";
 RT Mol. Cell. Biol. 10:4770-4777(1990).
 RU [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,

RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Cobby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.R.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grahame D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle B., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Leharasallo M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McKay K., McMurray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Patchlingam S.R., Plumb R.W., Ramsey H.,
 RA Rice C.W., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sultson J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Williams S.A.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RA "The DNA sequence and comparative analysis of human chromosome 20.";
 RA Nature 414:865-871(2001).
 [3]
 RX STRUCTURE BY NMR OF SH3 DOMAIN.
 RX MEDLINE=93208890; PubMed=7681365;
 RA Kohda D., Hatanaka H., Oda K., Mandlyan V., Ullrich A.,
 RA Schlesinger J., Inagaki F.;
 RT "Solution structure of the SH3 domain of phospholipase C-gamma.";
 RT Cell 72:953-960(1993).
 CC -1- FUNCTION: PLC-GAMMA IS A MAJOR SUBSTRATE FOR HEPARIN-BINDING
 CC GROWTH FACTOR 1 (ACIDIC FIBROBLAST GROWTH FACTOR)-ACTIVATED
 CC TYROSINE KINASE.
 CC -1- CATALYTIC ACTIVITY: 1-phosphatidy-1-D-myo-inositol 4,5-
 CC bisphosphate + H(2)O = D-myo-inositol 1,4,5-trisphosphate +
 CC diacylglycerol.
 CC -1- COFACTOR: Calcium.
 CC -1- PTM: THE RECEPTOR-MEDIATED ACTIVATION OF PLC-GAMMA 1 AND 2
 CC INVOLVES THEIR PHOSPHORYLATION BY TYROSINE KINASES IN RESPONSE
 CC TO LIGATION OF A VARIETY OF GROWTH FACTOR RECEPTORS AND IMMUNE
 CC SYSTEM RECEPTORS.
 CC -1- SIMILARITY: DOMAINS X AND Y ARE CONSERVED IN DIFFERENT FORMS
 CC OF PLC AND ARE ESSENTIAL FOR CATALYTIC ACTIVITY.
 CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 PH DOMAINS. THE SECOND ONE IS IN TWO
 CC PARTS.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 EF-HAND CALCIUM-BINDING DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M34667; AAA36452.1; -;
 DR EMBL; A022394; CAAL8537.1; -;
 DR PIR; A36466; A36466.
 DR PDB; 2HSP; 31-AUG-94.
 DR PDB; 1HSQ; 31-AUG-94.
 DR Genew; HGNC:9065; PLCG1.
 DR MIM; 172420;
 DR InterPro; IPR000008; C2.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR001192; PI_PLC.
 DR InterPro; IPR000909; PI_PLC_Xdom.
 DR InterPro; IPR001711; PI_PLC_Y.
 DR InterPro; IPR000980; SH2.
 DR InterPro; IPR001452; SH3.

DR Pfam; PF00017; SH2; 2.
 DR Pfam; PF00018; SH3; 1.
 DR Pfam; PF00168; C2; 1.
 DR Pfam; PF00169; PH; 2.
 DR Pfam; PF00387; PI-PLC-X; 1.
 DR Pfam; PF00388; PI-PLC-X; 1.
 DR PRINTS; PR00390; PPHPLIPASEC.
 DR PRINTS; PR00401; SH2DOMAIN.
 DR PRINTS; PR00452; SH3DOMAIN.
 DR ProDom; PD000066; SH3; 1.
 DR ProDom; PD000093; SH2; 2.
 DR ProDom; PD001202; PI_PLC_Y; 1.
 DR SMART; SM00239; C2; 1.
 DR SMART; SM00239; PH; 2.
 DR SMART; SM00148; PLCKC; 1.
 DR SMART; SM00149; PLCYC; 1.
 DR SMART; SM00252; SH2; 2.
 DR SMART; SM00326; SH3; 1.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS50001; SH2; 2.
 DR PROSITE; PS50002; SH3; 1.
 DR PROSITE; PS50003; C2 DOMAIN; 2.
 DR PROSITE; PS50004; C2 DOMAIN; 2.
 DR PROSITE; PS50007; PIPLC_X DOMAIN; 1.
 DR PROSITE; PS50008; PIPLC_Y DOMAIN; 1.
 KW Hydrolase; Lipid degradation; Transducer; SH2 domain; SH3 domain;
 Repeat; Calcium-binding; Phosphorylation; 3D-structure; Polymorphism.
 FT DOMAIN 27 142 PH 1.
 FT CA BIND 165 176 EF_HAND (POTENTIAL).
 FT DOMAIN 320 464 DOMAIN X.
 FT DOMAIN 489 523 PH 2 (FIRST PART).
 FT DOMAIN 550 657 SH2 1.
 FT DOMAIN 668 756 SH2 2.
 FT DOMAIN 791 851 SH3.
 FT DOMAIN 895 931 PH 2 (SECOND PART).
 FT DOMAIN 953 1070 DOMAIN Y.
 FT DOMAIN 1075 1177 C2 DOMAIN.
 FT ACT_SITE 1335 335 BY SIMILARITY.
 FT ACT_SITE 380 380 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 771 771 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 783 783 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 1253 1253 I -> T (IN DBSNP:753381).
 FT VARIANT 813 813 /FTID=VAR_011908.
 SQ SEQUENCE 1290 AA; 148531 MW; A805A8E2A18EDDAC CRC64;
 Query Match 15.4%; Score 62.5; DB 1; Length 1290;
 Best Local Similarity 28.1%; Pred. No. 19;
 Matches 18; Conservative 9; Mismatches 34; Indels 3; Gaps 1
 QY 5 VANNQGPNNRVKIKVSVNIIHGRIDIPASNPFLQ---KNTIGNGRRISLKLTSNGKI 61
 DB 883 IARRPEKKNRLRPFVFSISMAVAMSLDVAADSCSEFLQDWVKIKIRVAYQIADRLRIGKI 942
 QY 62 ASRR 65
 DB 943 MERR 946
 RESULT 12
 PAGE ECOLI
 ID PAPE ECOLI STANDARD; PRT; 173 AA.
 AC P08407;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Fimbrial protein pape precursor.
 GN PAPE.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OC NCBI_taxid=562;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-J96;
 RX MEDLINE=86149403; PubMed=2869489;
 RA Lindberg F., Lund B., Normark S.;
 RT "Gene products specifying adhesion of uropathogenic Escherichia coli
 RL are minor components of pili.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:1891-1895(1986).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN-J96;
 RC MEDLINE=88169520; PubMed=2895103;
 RX Lund B., Lindberg F., Normark S.;
 RA "Structure and antigenic properties of the tip-located P pilus
 RT proteins of uropathogenic Escherichia coli.";
 RL J. Bacteriol. 170:1887-1894(1988).
 RN [3]
 RN SEQUENCE FROM N.A.
 RP STRAIN-J96;
 RC MEDLINE=9302852; PubMed=1357526;
 RX Marklund B.-I., Tennent J.M., Garcia E., Hamers A., Baga M.,
 RA Lindberg F., Gaastria W., Normark S.;
 RT "Horizontal gene transfer of the Escherichia coli pap and prs pili
 RL operons as a mechanism for the development of tissue-specific
 RT adhesive properties.";
 RL Mol. Microbiol. 6:2225-2242(1992).
 CC -1- FUNCTION: FIMBRIAE (ALSO CALLED PILI), POLAR FILAMENTS RADIATING
 CC FROM THE SURFACE OF THE BACTERIUM TO A LENGTH OF 0.5-1.5
 CC MICROMETERS AND NUMBERING 100-300 PER CELL, ENABLE BACTERIA TO
 CC COLONIZE THE EPITHELIAL OF SPECIFIC HOST ORGANS.
 CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE MINOR COMPONENTS OF PILI.
 CC PILI WITH A DEFECTIVE PAPE GENE WILL HAVE LOW ADHESIVE CAPACITY
 CC OR NONE; HOWEVER, THE BINDING PROPERTY OF THE WHOLE CELL WILL NOT
 CC BE AFFECTED.
 CC -1- SUBCELLULAR LOCATION: SECRETED, TIP OF THE PILI.
 CC -1- DISEASE: STRAINS OF E. COLI THAT CAUSE INFECTION OF THE HUMAN
 CC URINARY TRACT PRODUCE PAP-PILI WHICH ARE HAIR-LIKE APPENDAGES
 CC CONSISTING OF ABOUT 1000 HELICALLY ARRANGED SUBUNITS OF THE
 CC PROTEIN PAPA. THESE PILI MEDIATE BINDING TO DIGALACTOSIDE-
 CC CONTAINING GLYCOLIPIDS PRESENT ON THE EPITHELIAL CELLS WHICH
 CC LINE THE URINARY TRACT.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X61238; CAA43556.1; -
 CC EMBL; M13239; AAA24280.1; -
 CC EMBL; M20146; AAA24288.1; -
 CC EMBL; X61239; CAA43568.1; -
 CC PIR; A25134; YOECEB
 CC InterPro; IPR000259; Fimbrial.
 CC InterPro; IPR004086; FimbrialPAPE.
 CC Pfam; PF00419; Fimbrial; 1.
 CC PRINTS; PR01555; FIMBRIALPAPE.
 CC Fimbrlia; Cell adhesion; Signal.
 CC SIGNAL 1 24 POTENTIAL.
 CC CHAIN 25 173 FIMBRIAL PROTEIN PAPE.
 CC SQUENCE 173 AA; 18569 MW; E27577D09C46A863 CRC64;
 CC
 CC Query Match 15.3%; Score 62; DB 1; Length 173;
 CC Best Local Similarity 27.0%; Pred. No. 2,4;
 CC Matches 20; Conservative 11; Mismatches 35; Indels 8; Gaps 3
 CC
 CC 1 MKTVV-AFNOGPNRRVYIKVSNVLIHGRKIDIPASKNPHLQKNTIGCRISLKTSG 59
 CC Db MKVTITATNTNY--NNALIVONTSTSDGLVLYLNS-----NAGNIGTATITGPTPTG 131
 CC 60 KIAARVDGIIMAY 73

Db . 132 KITGNNADKTI SLH 145

ID	TOP1 THEME	STANDARD;	PRT;	633 AA.
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	DNA topoisomerase I (EC 5.99.1.2) (Omega-protein) (relaxing enzyme)			
DE	(Unwisting enzyme) (Switzerland).			
OS	TOPA OR TM0258			
OC	Thermotoga maritima.			
OC	Bacteria; Thermotogae; Thermotogae (class); Thermotogales;			
OC	Thermotogaceae; Thermotoga.			
OX	NCBI_TaxID=2336;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MSB / DSM 3109;			
RC	MEDLINE=99287316; PubMed=10360571;			
RA	Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,			
RA	Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,			
RA	McDonald L., Utecher B.T., Malek J.A., Liner K.D., Garrett M.M.,			
RA	Stewart A.M., Sutton M.D., Pratt M.S., Phillips C.A., Richardson D.,			
RA	Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,			
RA	Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;			
RT	"Evidence for lateral gene transfer between Archaea and Bacteria from			
RT	genome sequence of Thermotoga maritima.";			
RL	Nature 399:323-329 (1999).			
CC	-1- FUNCTION: THE REACTION CATALYZED BY TOPOISOMERASE LEADS TO THE			
CC	CONVERSION OF ONE TOPOLOGICAL ISOMER OF DNA TO ANOTHER.			
CC	-1- CATALYTIC ACTIVITY: ATP-independent breakage of single-stranded			
CC	DNA, followed by passage and rejoining.			
CC	-1- SUBUNIT: MONOMER (BY SIMILARITY).			
CC	-1- MISCELLANEOUS: WHEN A TOPOISOMERASE TRANSIENTLY BREAKS A DNA			
CC	BACKBONE BOND, IT SIMULTANEOUSLY FORMS A PROTEIN-DNA LINK, IN			
CC	WHICH A TYROSYL OXYGEN IN THE ENZYME IS JOINED TO A DNA PHOSPHORUS			
CC	AT ONE END OF THE ENZYME-SEVERED DNA STRAND.			
CC	-1- SIMILARITY: BELONGS TO PROKARYOTIC TYPE I/III TOPOISOMERASE			
CC	FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; U27841; AAA68949.1; -			
DR	EMBL; AE001708; AAD35346.1; -			
DR	HSSP; P06612; 1ECL.			
DR	TIGR; TM0258; -			
DR	InterPro; IPR002936; DNAPrim_toprim.			
DR	InterPro; IPR003601; DNATopi_ATP_bind.			
DR	InterPro; IPR003602; DNATopi_DNA_bind.			
DR	InterPro; IPR000380; Prok_Epiomase.			
DR	Pfam; PF01131; Topoisom_bac; 1.			
DR	Pfam; PF01751; Toprim; 1.			
DR	PRINTS; PRO0417; PRPISMRASE1.			
DR	SMART; SM00437; TOP1ac; 1.			
DR	SMART; SM00436; TOP1bc; 1.			
DR	SMART; SM00493; TOP1rim; 1.			

DR	TIGRFAMs: TIGR01051; topA_bact; 1
DR	PROSITE: PS00396; TOPOLISOMERASE 1_PROK; 1.
KW	Isomerase, Topoisomerase; DNA-binding; Zinc-finger; Metal-binding;
KW	Complete proteome.
FT	ZN_FING 559
FT	C4-TYPE. 580
FT	ACT_SITE 288
SC	SEQUENCE 633 AA; 72694 MW; F7262A04060CFE9 CRC64;
Query March 15.1%; Score 61.5; DB 1; Length 633;	
Best Local Similarity 25.4%; Pred. No. 12;	
Matches 16; Conservative 14; Mismatches 32; Indels 1; Gaps 1	
OY	18 IKRVSVIIHGRRIDIPASKNFHLQKNTIGTCRRISLTGTSGIASRRVGGIIAAVQNP 77
Db	96 IARVYNTLGRKRNRI-VFSRITPRVIRBAVNKPREDIMKVTQAQGLAKRIILRIVIGYSLSPV 154
OY	78 SWK 80
Db	155 LMR 157

ID	RS18 DROME	STANDARD:	PRT:	152 AA.
AC	P41094, Q9V911;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	40S ribosomal protein S18.			
GN	RP518 OR CG8900.			
OS	Drosophila melanogaster (fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;			
OC	Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;			
OC	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
CC	NCBI_TaxId=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A. (LONG ISOFORM).			
RC	STRAIN=Canton-S;			
RA	MEDLINE=944215909; PubMed=8163194;			
RX	Garwood J., Lepesant J.-A.;			
RT	"The Drosophila melanogaster homolog of ribosomal protein S18.";			
RL	Gene 141:231-235(1994).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Berkeley;			
RX	MEDLINE=20196005; PubMed=10731122;			
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wotman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,			
RA	Abiril J.F., Agbayan A., An H.-D., Andrews-Pfankuch C., Baldwin D.,			
RA	Ballew R.M., Baau A., Bakendle J., Bayraktaroglu L., Beasley E.M.,			
RA	Beehon K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Borchan M.R., Bouck J., Brokstein P., Brothier P.,			
RA	Buttis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cateley S., Dahlke C., Davenport L.B., Davies P.,			
RA	DePalos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Fowler C., Gabriellian A.E., Gang N.S., Gelbart W.M., Glaeser K.,			
RA	Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Howland T.J., Hernandez J.R., Houck J.,			
RA	Hoskins D., Houston K.A., Houtman T.J., Wei M.-H., Ibegwam C.,			
RA	Jajalil B., Kalish F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kammel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			
RA	Liu X., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,			
RA	Lesko P., Lei Y., McIntosh T.C., McLeod M.P., McPherson D.,			
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,			
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,			
RA	Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paaleb J.M.,			
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,			

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Sylvestre R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Glibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [3]
 RP REVISIONS, AND ALTERNATIVE SPLICING.
 RC STRAIN-BEKELEY;
 RA Maira S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochman S.E., Smith C.D.,
 RA Tupy J.L., Bergman C.M., Berman B.P., Carlson J.W., Celisner S.E.,
 RA Clamp M.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,
 RA Harris N.L., Kronmiller B., Marshall B., Millburn G.H., Richter J.,
 RA Russo S., Searle S.M.J., Smith E., Shu S., Smutnak F.,
 RA Whitefield E.O., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Mungall C.J., Lewis S.E.;
 RT "Annotation of *Drosophila melanogaster* genome.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: INVOLVED IN THE BINDING OF FMET-TRNA AND, HENCE, IN THE
 CC INITIATION OF TRANSLATION (BY SIMILARITY).
 CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
 CC short form; are produced by alternative splicing.
 CC -!- SIMILARITY: BELONGS TO THE S13P FAMILY OF RIBOSOMAL PROTEINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL, L22959; AAA28870.1; -;
 DR EMBL, AE003792; AAM68401.1; -;
 DR EMBL, AE003792; AAM68402.1; -;
 DR Flybase; FBgn0010411; RPS18.
 DR InterPro; IPR001892; Ribosomal_S13.
 DR Pfam; PF00416; Ribosomal_S13; 1.
 DR Prodom; PD001363; Ribosomal_S13; 1.
 DR PROSITE; PS00646; RIBOSOMAL_S13; 1.
 DR Ribosomal protein; Alternative splicing.
 KM VARSPPLIC 58
 FT ECTEEVDKVTITISNPLOKYPNNFELNROKDIIDGKYWL
 FT TSS -> RQGGDHLPSAVQAGQVLVPOAGGHRMQLAA
 FT DILQGLENA (IN SHORT ISOFORM).
 FT MISSING (IN SHORT ISOFORM).
 FT VARSPPLIC 102 152
 FT SEQUENCE 152 AA; 17611 MW; F52P9B865B880EF CRC64;
 SQ
 Query Match 15.0%; Score 61; DB 1; Length 152;
 Best Local Similarity 25.6%; Pred. No. 2.8; Indels 22; Gaps 3;
 Matches 21; Conservative 10; Mismatches 29;
 QY 12 PNRPRVFIKRVSYNIHGR---IDIFASKNFHLQKNTIGTGR-----ISIKLTS 58
 DB 6 PEKFGHILKIMNTNIDKRVKGIAMTAIK-----GVRKYNSIVLKKADVDTKRA 56
 QY 59 GKIASRRVDGIIAAYONPASMK 80
 DB 57 GECTEEVDKVTITISNPLOK 78
 RESULT 15
 ID Y12K_SMSV4 STANDARD; PRT; 110 AA.
 AC P36289;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Possible 12 kDa nucleic acid-binding protein.
 OS San Miguel sea lion virus (serotype 4) (SMSV 4).
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 CC Vestivirus
 OK NCBI_TaxID=36407;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92410750; PubMed=1529644;
 RA Neill J.D.;
 RT "Nucleotide sequence of the capsid protein gene of two serotypes of
 RT San Miguel sea lion virus: identification of conserved and non-
 RT conserved amino acid sequences among calicivirus capsid proteins.";
 RL Virus Res. 24:211-222(1992).
 CC -!- SIMILARITY: TO FELINE CALICIVIRUS 12 KDA PROTEIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL, M87482; AAA16221.1; -;
 DR PIR; D48562; D48562.
 DR SEQUENCE 110 AA; 12566 MW; 1425SD593827418E CRC64;
 SQ
 Query Match 14.9%; Score 60.5; DB 1; Length 110;
 Best Local Similarity 32.4%; Pred. No. 2.2;
 Matches 11; Conservative 10; Mismatches 12; Indels 1; Gaps 1;
 QY 17 FIKRVSNVTHHGRDIDIFASKNFHLQKNTIGTGR 50
 DB 10 FLNSVANAVEGKLDL-ASKGLQKSRALDTER 42

Search completed: April 20, 2003, 13:07:25
 Job time : 17.5789 secs

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:57:30 ; Search time 32.4211 Seconds

(without alignments)
237.215 Million cell updates/sec

Title: US-09-142-524D-1

Perfect score: 406

Sequence: 1 MKYTVAFNPGFNRVFIKR.....IASRRVGIITAYQNPASWK 80

Scoring table: BLASTUM62

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99	24.4	514	2 S48730	Cry j II protein -
2	99	24.4	514	2 JC2498	second major aller
3	74	18.2	514	2 JC7100	polygalacturonase
4	71	17.5	374	2 JC2124	major allergen Cry
5	71	17.5	374	2 JC2123	Jun a 2 protein -
6	71	17.5	507	2 UC7366	ribosomal protein
7	70	17.2	152	1 S30393	pectate lyase like
8	70	17.2	152	1 R3R118	ABC transporter (A
9	70	17.2	152	1 R3R118	pectate lyase like
10	68	16.7	394	2 T49115	ABC transporter (A
11	67	16.5	674	2 F83794	pectate lyase like
12	66.5	16.4	196	2 S54580	pectate lyase like
13	66	16.3	394	2 S75065	pectate lyase like
14	66	16.3	394	2 T49116	pectate lyase like
15	66	16.3	464	2 F89828	pectate lyase like
16	65.5	16.1	391	2 F72552	pectate lyase like
17	65.5	16.1	531	2 T33980	pectate lyase like
18	65	16.0	390	2 H86253	pectate lyase like
19	65	16.0	498	2 UC3353	pectate lyase like
20	64.5	15.9	770	2 AH2465	pectate lyase like
21	64	15.8	222	2 F72335	pectate lyase like
22	64	15.8	398	2 F82298	pectate lyase like
23	63.5	15.6	433	2 T34458	pectate lyase like
24	63.5	15.6	533	2 T34458	pectate lyase like
25	62.5	15.4	629	2 T1776	pectate lyase like
26	62.5	15.4	1290	2 A36466	pectate lyase like
27	62	15.3	173	1 Y0ECP6	pectate lyase like
28	62	15.3	368	2 A90266	pectate lyase like
29	62	15.3	368	2 A90266	pectate lyase like

30	62	15.3	5627	2 C83339	hypothetical prote
31	61.5	15.1	369	2 E75620	hypothetical prote
32	61.5	15.1	633	2 S62737	DNA topoisomerase
33	61	15.0	173	2 E27743	pape fibrillar prot
34	60.5	14.9	110	2 D48562	hypothetical nucle
35	60.5	14.9	856	2 B81399	hypothetical prote
36	60	14.8	154	2 T27228	probable periplasm
37	60	14.8	219	2 C84647	ribosomal protein
38	60	14.8	275	2 T28738	probable synaptoch
39	60	14.8	295	2 S60711	hypothetical prote
40	60	14.8	410	2 H90304	band-6-protein - b
41	60	14.8	627	2 T00484	conserved hypochet
42	60	14.8	725	2 S60712	hypothetical prote
43	60	14.8	844	1 TVMSV	band-6-protein - h
44	59.5	14.7	254	2 D69545	transforming prote
45	59.5	14.7	263	2 A72349	ABC transporter, A
					conserved hypochet

ALIGNMENTS

RESULT 1
S48730
Cry j II protein - Japanese cedar
C:Species: Cryptomeria japonica (Japanese cedar)
C:Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 28-May-1999
A:Accession: S48730
R:Nanda, M.; Toriige, K.; Hino, K.; Taniguchi, Y.; Fukuda, S.; Usui, M.; Kur
FBS Lett. 353, 124-128, 1994
A:Title: Molecular cloning of the second major allergen, Cry j II, from Japanese cedar p
A:Reference number: S48730; MUID:95010777; PMID:7926035
A:Accession: S48730
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-514 <NM>
A:Cross-references: GB:D37765; NID:G577695; PIDN:BA07021.1; PID:dl007598; PID:G577696

Query Match 24.4%; Score 99; DB 2; Length 514;
Best Local Similarity 100.0%; Pred. No. 0.00095;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 IDIFASKNFHLQKNTIGTG 49
DB 236 IDIFASKNFHLQKNTIGTG 254

RESULT 2
JC2498
second major allergen Cry j II precursor - Japanese cedar
C:Species: Cryptomeria japonica (Japanese cedar)
C:Date: 16-Mar-1995 #sequence_revision 26-May-1995 #text_change 21-Jul-2000
A:Accession: JC2498; PC2346; A60147
R:Komiyama, N.; Sone, T.; Shimizu, K.; Morikubo, K.; Kino, K.
Biochem. Biophys. Res. Commun. 201, 1021-1028, 1994
A:Title: cDNA cloning and expression of Cry j II, the second major allergen of Japanese
A:Reference number: JC2498; MUID:94271186; PMID:8002972
A:Accession: JC2498
A:Molecule type: mRNA
A:Residues: 1-514 <NM>
A:Cross-references: DDBJ:D29772; NID:G506857; PIDN:BA06172.1; PID:G506858
A:Accession: PC2346
A:Molecule type: protein
A:Residues: 52-61 <KO2>
R:Sakaguchi, M.; Inouye, S.; Tanai, M.; Ando, S.; Usui, M.; Matuhasi, T.
Allergy 45, 309-312, 1990
A:Title: Identification of the second major allergen of Japanese cedar pollen.
A:Reference number: A60147; MUID:90342988; PMID:2382797
A:Accession: A60147
A:Molecule type: protein
A:Residues: 55-64 <SAK>
C:Keywords: glycoprotein; pollen
F:1-54/Domain: signal sequence #status predicted <SIG>
F:55-460/Product: second major allergen Cry j #status predicted <MAT>

F:429,460,472/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 24.4%; Score 99; DB 2; Length 514;
Best Local Similarity 100.0%; Pred. No. 0.00055;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 IDIFASKNFHLOKNTIGTG 49
|||||
DB 236 IDIFASKNFHLOKNTIGTG 254

RESULT 3

JC7100
polygalacturonase Cha o 2 - Japanese cypress
C/Species: Chamaecyparis obtusa (Japanese cypress)
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
C/Accession: J07100; PC7026
R/Mori, T.; Yokoyama, M.; Komiyama, N.; Okano, M.; Kino, K.
Biochem. Biophys. Res. Commun. 263, 166-171, 1999
A/Title: Purification, identification, and cDNA cloning of Cha o 2, the second major all
A/Reference number: J07100; MUID:99417540; PMID:10486272
A/Accession: J07100
A/Molecule type: protein
A/Residues: 1-514 <MOR>
A/Accession: PC7026
A/Molecule type: protein
A/Residues: 51-62 <MO2>

Query Match 18.2%; Score 74; DB 2; Length 514;
Best Local Similarity 73.7%; Pred. No. 0.87;
Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 31 IDIFASKNFHLOKNTIGTG 49
|||||
DB 236 IDIFASKNFHLOKNTIGTG 254

RESULT 4

JC2124
major allergen Cry j I precursor (clone pCCI-15) - Japanese cedar
C/Species: Cryptomeria japonica (Japanese cedar)
C/Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 21-Jul-2000
C/Accession: J02124
R/Sone, T.; Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.
Biochem. Biophys. Res. Commun. 199, 619-625, 1994
A/Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japanese
A/Reference number: J02123; MUID:94183224; PMID:8135802
A/Accession: J02124
A/Molecule type: mRNA
A/Residues: 1-374 <SON>
A/Cross-references: GB:D26545; NID:9493633; PIDN:BA05543.1; PID:9493634
A/Experimental source: pollen
A/Note: the authors described carbohydrate binding site for residue 279
C/Superfamily: pectate lyase LAR59
C/Keywords: glycoprotein; pollen
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-374/Product: major allergen Cry j I (clone pCCI-15) #status predicted <MAT>
F:158,191,293,354/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 17.5%; Score 71; DB 2; Length 374;
Best Local Similarity 55.6%; Pred. No. 1.4;
Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKVTVAFNQGPGRRVPIKRVSNVITH 27
|||||
DB 234 MKVTVAFNQGPGRRVPIKRVSNVITH 260

RESULT 5

JC2123
major allergen Cry j I precursor (clone pCCI-2-2) - Japanese cedar
C/Species: Cryptomeria japonica (Japanese cedar)
C/Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 21-Jul-2000

C/Accession: J02123; PC2065
R/Sone, T.; Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.

Biochem. Biophys. Res. Commun. 199, 619-625, 1994
A/Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japanese
A/Reference number: J02123; MUID:94183234; PMID:8135802
A/Accession: J02123

A/Molecule type: mRNA
A/Residues: 1-374 <SON>
A/Cross-references: GB:D26544; NID:9493631; PIDN:BA05542.1; PID:9493632
A/Experimental source: pollen
A/Accession: PC2065

A/Molecule type: protein
A/Residues: 22-53,58-81,219-232,236-258,299-307,346-372 <SO2>
A/Note: the authors described carbohydrate binding site for residue 279
C/Superfamily: pectate lyase LAR59
C/Keywords: glycoprotein; pollen

F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-374/Product: major allergen Cry j I (clone pCCI-2-2) #status predicted <MAT>
F:158,191,293,354/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 17.5%; Score 71; DB 2; Length 374;
Best Local Similarity 55.6%; Pred. No. 1.4;
Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKVTVAFNQGPGRRVPIKRVSNVITH 27
|||||
DB 234 MKVTVAFNQGPGRRVPIKRVSNVITH 260

RESULT 6

J07366
Jun a 2 protein - mountain cedar
C/Species: Juniperus ashei (mountain cedar)
C/Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000
C/Accession: J07366; PC7093
R/Yokoyama, M.; Miyahara, M.; Shimizu, K.; Kino, K.; Tsunoo, H.
Biochem. Biophys. Res. Commun. 275, 195-202, 2000
A/Title: Purification, identification, and cDNA cloning of Jun a 2, the second major all
A/Reference number: J07366
A/Accession: J07366
A/Molecule type: mRNA
A/Residues: 1-507 <YOK>
A/Cross-references: GB:AJ404653
A/Accession: PC7093
A/Molecule type: protein
A/Residues: 55-63 <YOK>
A/Comment: This protein, a second major allergen of mountain cedar pollen, which is invo
o the polygalacturonase family.
C/Keywords: glycoprotein; pollen

Query Match 17.5%; Score 71; DB 2; Length 507;
Best Local Similarity 73.7%; Pred. No. 1.9;
Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 31 IDIFASKNFHLOKNTIGTG 49
|||||
DB 237 IDIFASKNFHLOKNTIGTG 255

RESULT 7

S30393
ribosomal protein S18, cytosolic - human
C/Species: Homo sapiens (man)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C/Accession: S30393; S68935
R/Chassin, D.; Bellec, D.; Koman, A.
Nucleic Acids Res. 21, 745, 1993
A/Title: The human homolog of ribosomal protein S18.
A/Reference number: S30393; MUID:93181276; PMID:8441687
A/Accession: S30393

A/Molecule type: mRNA
A/Residues: 1-152 <CHA>
A/Cross-references: EMBL:X69150; NID:938422; PIDN:CA856794.1; PID:96006558

R.Vladimirov, S.N.; Ivanov, A.V.; Karpova, G.G.; Musolyanov, A.K.; Egorov, T.A.; Thiede, E. J. Biochem. 239, 144-149, 1996
 A>Title: Characterization of the human small ribosomal subunit proteins by N-terminal an
 A:Reference number: S68911; MUID:96305378; PMID:8706659
 A:Accession: S68935

A:Molecule type: protein
 A:Residues: 5', 56-57, 'X', 59-64, 'XX', 67-69 <VLA>
 C:Genetics:

A:Gene: GDB:RPS18
 A:Cross-references: GDB:138447; OMIM:180473
 A:Map position: 6p21.3-6p21.3

C:Superfamily: Escherichia coli ribosomal protein S13
 C:Keywords: blocked amino end; protein biosynthesis; ribosome

Query Match 17.2%; Score 70; DB 1; Length 152;
 Best Local Similarity 30.4%; Pred. No. 0.68;

Matches 24; Conservative 7; Mismatches 32; Indels 16; Gaps 2;

QY 12 PNRRVFKRVSNVTHGRIDIFASKNFHLQKNTGTGR-----ISLKLTSGKI 61
 DB 6 PEKQHLRLVLTNTIDGRKIAFAITAK-----GVGRYAHVVLKADIDLTGRAGEL 59

QY 62 ASRRVDGIILAYONPASMK 80
 DB 60 TEDEVERVITIMONPROYK 78

RESULT 8

R3RT18
 ribosomal protein S18, cytosolic [validated] - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Jul-2000

C:Accession: JH0419
 R:Chan, Y.L.; Paz, V.; Wool, I.G.
 Biochem. Biophys. Res. Commun. 178, 1212-1218, 1991

A>Title: The primary structure of rat ribosomal protein S18.
 A:Reference number: JH0419; MUID:91337062; PMID:1872840
 A:Accession: JH0419

A:Molecule type: mRNA
 A:Residues: 1-152 <CHA>
 A:Cross-references: GB:X57529; NID:9433446; PIDN:CAA0750.1; PID:9433447

A>Note: the protein is designated as ribosomal protein S18 according to conigration anal
 C:Superfamily: Escherichia coli ribosomal protein S13
 C:Keyword: protein biosynthesis; ribosome

Query Match 17.2%; Score 70; DB 1; Length 152;
 Best Local Similarity 30.4%; Pred. No. 0.68;

Matches 24; Conservative 7; Mismatches 32; Indels 16; Gaps 2;

QY 12 PNRRVFKRVSNVTHGRIDIFASKNFHLQKNTGTGR-----ISLKLTSGKI 61
 DB 6 PEKQHLRLVLTNTIDGRKIAFAITAK-----GVGRYAHVVLKADIDLTGRAGEL 59

QY 62 ASRRVDGIILAYONPASMK 80
 DB 60 TEDEVERVITIMONPROYK 78

RESULT 9

176666
 ribosomal protein S18 [similarity] - mouse
 N:Alternate names: ribosomal protein S13 [misnomer]

C:Species: Mus musculus (house mouse)
 C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 03-Nov-2000
 C:Accession: I76666; I57006

R:MacMurray, A.J.; Shin, H.S.
 Mamm. Genome 2, 87-95, 1992
 A>Title: The murine hMC encodes a mammalian homolog of bacterial ribosomal protein S13.

A:Reference number: I57006; MUID:92182530; PMID:1543907

A:Accession: I76666
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-152 <RES>

A:Cross-references: GB:M76763; NID:g198579; PIDN:AAA16795.1; PID:g198580

A:Accession: I57006
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 1-103, 'H', 105-152 <RE2>
 A:Cross-references: GB:M76762; NID:g198577; PIDN:AAA16796.1; PID:g198578

C:Genetics:
 A:Gene: Ke-3
 A:Introns: 1/3, 34/3, 97/3, 128/2

C:Superfamily: Escherichia coli ribosomal protein S13

Query Match 17.2%; Score 70; DB 2; Length 152;
 Best Local Similarity 30.4%; Pred. No. 0.68;

Matches 24; Conservative 7; Mismatches 32; Indels 16; Gaps 2;

QY 12 PNRRVFKRVSNVTHGRIDIFASKNFHLQKNTGTGR-----ISLKLTSGKI 61
 DB 6 PEKQHLRLVLTNTIDGRKIAFAITAK-----GVGRYAHVVLKADIDLTGRAGEL 59

QY 62 ASRRVDGIILAYONPASMK 80
 DB 60 TEDEVERVITIMONPROYK 78

RESULT 10

T49115
 pectate lyase like protein - Arabidopsis thaliana
 N:Alternate names: protein AT4g22080
 C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Sep-2000
 C:Accession: T49115
 R:Bevan, M.; Medler, H.; Mambutt, R.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.; M

submitted to the Protein Sequence Database, May 2000
 A:Reference number: Z25016
 A:Accession: T49115

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-394 <BRV>

A:Cross-references: EMBL:AL022140; GSPDB:GNO0062; ATSP:AT4g22080
 A:Experimental source: cultivar Columbia; BAC clone FLN20
 C:Genetics:
 A:Gene: ATSP:AT4g22080
 A:Map position: 4
 A:Introns: 240/3; 275/2; 321/3

C:Superfamily: pectate lyase L4T59

Query Match 16.7%; Score 68; DB 2; Length 394;
 Best Local Similarity 28.6%; Pred. No. 3.3;

Matches 24; Conservative 7; Mismatches 27; Indels 26; Gaps 3;

QY 1 MKTVAFNQGPNRRVFKRVSNVTH--GRRID-----IFASKNFHLQKN 44
 DB 256 MKTVAFNHFQGLVERMPVRGRYAHVANNRXYDKIMYALGSSADPTIFSEGNVFIAD 315

QY 45 TIGTRRISLKLTSKGKIASRRVDG 68
 DB 316 -----KSNKSKYTKREYKNG 329

RESULT 11

F83794
 ABC transporter (ATP-binding protein) BH1158 [imported] - Bacillus halodurans (strain C-1)

C:Species: Bacillus halodurans
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 17-May-2002
 C:Accession: F83794

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiran
 Nucleic Acids Res. 28, 4317-4331, 2000
 A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

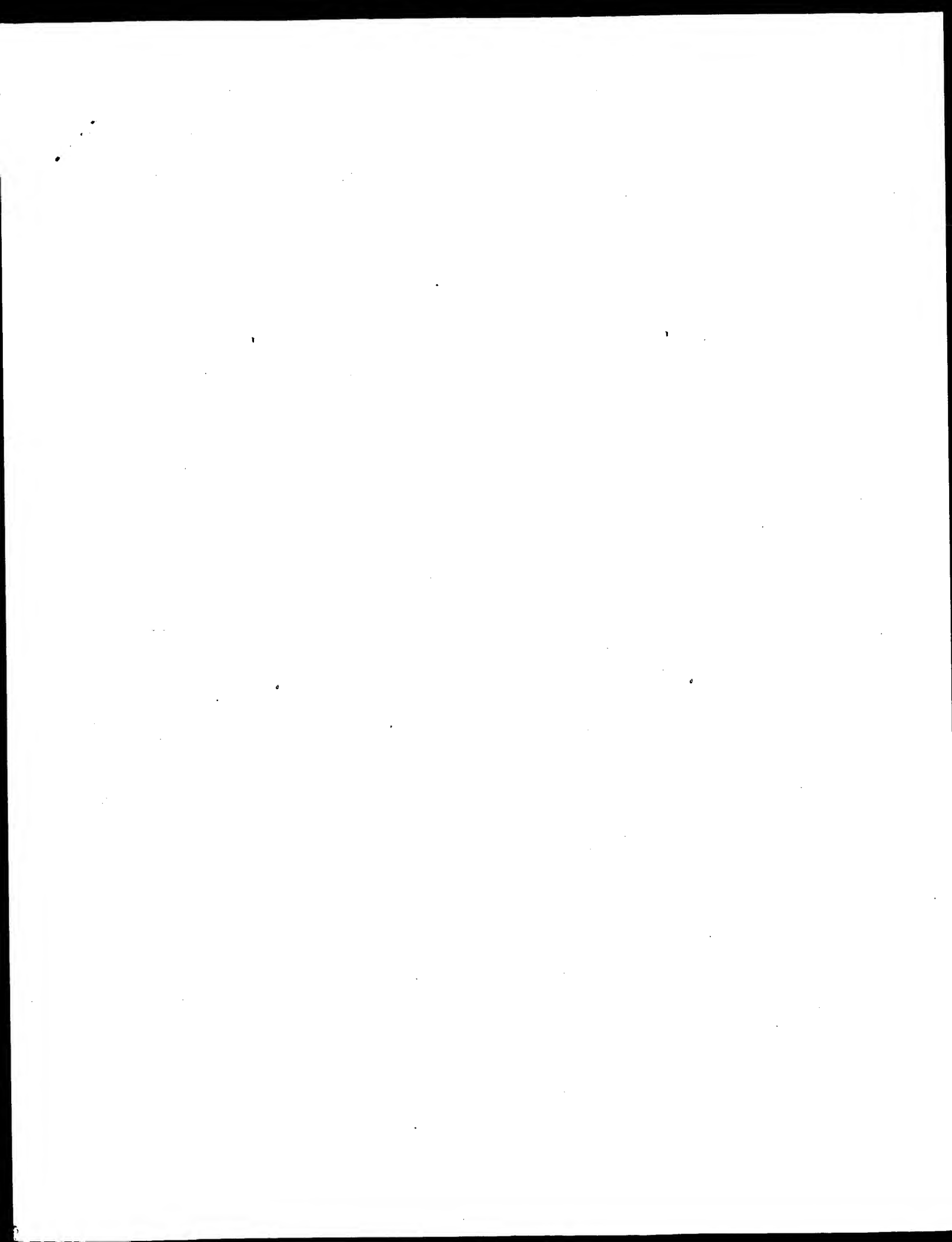
A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: F83794
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-674 <STO>

Query Match 16.3%; Score 66; DB 2; Length 464;
Best Local Similarity 28.6%; Pred. No. 6.9;
Matches 26; Conservative 9; Mismatches 28; Gaps 5

Qy 11 GPNRRVFIRVS-NVLI-HGRRIDIFASKNPHLOKNTIGTGRH----- 52
Db 302 GPNRFVEVKRTSGNVFIENGSK---FAKGGKYLKGKVAAGVGFQGMDDLANDNKTFGEA 358
Qy 53 ----SLKLTSGKIASBRVDGIIAAYQNPASW 79
Db 359 LSHNGMTLLAAGSVAA---FGVAALVSNPVGW 386

Search completed: April 20, 2003, 13:15:34
Job time : 35.4211 secs



X 10 X
VDGIIAAYONPASWK
| | | | | | | | | |
80 90 100 X
RLKMPYIAGYKTFDGRVVDGIIAAYONPASWK

3. US-09-142-524D-97 (1-15)
US-09-142-524D-1 Sequence 1, Application US/09142524D

Sequence 1, Application US/09142524D
GENERAL INFORMATION:

APPLICANT: Sone, Toshio
APPLICANT: Kume, Akimori
APPLICANT: Dairiki, Kazuo
APPLICANT: Iwama, Akiko
APPLICANT: Kino, Kohsuke
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease
FILE REFERENCE: SPO-103
CURRENT APPLICATION NUMBER: US/09/142,524D
CURRENT FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: PCT/JP97/00740
PRIOR FILING DATE: 1997-03-10
NUMBER OF SEQ ID NOS: 174
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 1
LENGTH: 80
TYPE: PRT
ORGANISM: Cryptomeria japonica

Initial Score = 15 Optimized Score = 15 Significance = 0.00
Residue Identity = 100% Matches = 15 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

X
VDGIIAA
| | | | |
MKYTVAFNQGPNRRVFIKRVSNVIHGRIDIFASKNPHLOKNTIGTGRRISLKTSGKIASRRVDGIIAA
10 20 30 40 50 60 70 X

10 X
YONPASWK
| | | | |
YONPASWK
80

> O <
01 0 IntellGenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-09-142-524d-100.res made by jdelaval on Wed 23 Apr 103 17:48:12-PDT.

Query sequence being compared: US-09-142-524D-100 (1-15)
Number of sequences searched: 3
Number of scores above cutoff: 3

Results of the initial comparison of US-09-142-524D-100 (1-15) with:
File : 09142524.pep

```

100-
N -
U -
M -
B -
E -
R -
O -
F -
S -
E -
Q -
U -
N -
M -
C -
E -
S -
SCORE 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
STDDEV -1

```

PARAMETERS

Similarity matrix Unitary 1 K-tuple 2
Mismatch penalty 1 Joining penalty 20
Gap penalty 1.00 Window size 15
Gap size penalty 0.05
Cutoff score 0
Randomization group 0

SEARCH STATISTICS

Scores: Mean 11 Median 4 Standard Deviation 6.93

Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00

Number of residues: 319
Number of sequences searched: 3
Number of scores above cutoff: 3

The scores below are sorted by initial score.
Significance is calculated based on initial score.

2 100% similar sequences to the query sequence were found:

Sequence Name Description Init. Opt. Length Score Score Sig. Frame

The list of other best scores is:

Sequence Name	Description	Length	Score	Opt.	Sig.	Frame
1. US-09-142-524D-3 Sequence 3, Application U		134	15	15	0.58	0
2. US-09-142-524D-2 Sequence 2, Application U		105	15	15	0.58	0
3. US-09-142-524D-1 Sequence 1, Application U		80	3	5	-1.15	0

1. US-09-142-524D-100 (1-15)
US-09-142-524D-3 Sequence 3, Application US/09142524D

Sequence 3, Application US/09142524D
GENERAL INFORMATION:

APPLICANT: Sone, Toshio
APPLICANT: Kume, Akimori
APPLICANT: Dairiki, Kazuo
APPLICANT: Iwama, Akiko
APPLICANT: Kino, Kohsuke
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Diseases
FILE REFERENCE: SPO-103
CURRENT APPLICATION NUMBER: US/09/142,524D
CURRENT FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: PCT/JP97/00740
PRIOR FILING DATE: 1997-03-10
NUMBER OF SEQ ID NOS: 174
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 134
TYPE: PRT
ORGANISM: Cryptomeria japonica

Initial Score = 15 Optimized Score = 15 Significance = 0.58
Residue Identity = 100% Matches = 15 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

```

MKTVAFNGPGRNRFVIFKRVSVNIIHGRRIDIFASKNFHLOKNTTGTGRMKNRRIWLFQAKLTGFTLMGR
10 20 30 40 50 60 70
RPLMIIFSGNNIKLMPMYIAGYKTFDGRBRAEVSYVHVGAKFIRVDGIIAAYONPASMK
80 90 100 110 120 130

```

2. US-09-142-524D-100 (1-15)
US-09-142-524D-2 Sequence 2, Application US/09142524D

Sequence 2, Application US/09142524D
GENERAL INFORMATION:

APPLICANT: Sone, Toshio
APPLICANT: Kume, Akimori
APPLICANT: Dairiki, Kazuo
APPLICANT: Iwama, Akiko
APPLICANT: Kino, Kohsuke
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Diseases
FILE REFERENCE: SPO-103
CURRENT APPLICATION NUMBER: US/09/142,524D
CURRENT FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: PCT/JP97/00740
PRIOR FILING DATE: 1997-03-10
NUMBER OF SEQ ID NOS: 174
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 105
TYPE: PRT
ORGANISM: Cryptomeria japonica

Initial Score = 15 Optimized Score = 15 Significance = 0.58
 Residue Identity = 100% Matches = 15 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X
 10 X
 NNRIMQFAKLTGFT
 |||||
 10
 MKVTVAFNQFGNRRVFIKRVSNVLIHGRIDIFASKNPHLOKNTIGTRKRWKNNRIWLPFAKLTGFTLMGR
 10 20 30 40 50 60 70
 80 90 100
 RLKMPMYIAGYKTFDGRRVGIIAAYQNPASWK

3. US-09-142-524D-100 (1-15)
 US-09-142-524D-1 Sequence 1, Application US/09142524D

Sequence 1, Application US/09142524D
 GENERAL INFORMATION:

APPLICANT: Sone, Toshio
 APPLICANT: Kume, Akinori
 APPLICANT: Daitiki, Kazuo
 APPLICANT: Iwama, Akiko
 APPLICANT: Kino, Kohsuke
 TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Diseases
 FILE REFERENCE: SPO-103
 CURRENT APPLICATION NUMBER: US/09/142,524D
 CURRENT FILING DATE: 1998-09-09
 PRIOR APPLICATION NUMBER: PCT/JP97/00740
 PRIOR FILING DATE: 1997-03-10
 NUMBER OF SEQ ID NOS: 174
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 1
 LENGTH: 80

TYPE: PRT
 ORGANISM: Cryptosporidia japonica

Initial Score = 3 Optimized Score = 5 Significance = -1.15
 Residue Identity = 40% Matches = 6 Mismatches = 6
 Gaps = 3 Conservative Substitutions = 0

X
 10 X
 NNRIMQFAKLTGFT
 |||||
 10
 MKVTVAFNQFGNRRVFIKRVSNVLIHGRIDIFASKNPHLOKNTIGTRRISL---KLTSGKIASRRVDCI
 10 20 30 40 50 60
 70
 IIAAYQNPASWK

Initial Score = 13 Optimized Score = 13 Significance = 0.69

> O < Intelligenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-09-142-524d-120.res made by jdelaval on Wed 23 Apr 103 17:48:52-PDT.

Query sequence being compared: US-09-142-524D-120 (1-15)
Number of sequences searched: 3
Number of scores above cutoff: 3

Results of the initial comparison of US-09-142-524D-120 (1-15) with:
File: 09142524.pep

```

100-
N  -
U  50-
M  -
B  -
E  -
R  -
O  -
F  10-
S  -
E  5-
Q  -
U  -
E  -
N  -
C  -
S  -
SCORE 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14
STDDEV

```

PARAMETERS

Similarity matrix Unitary K-tuple 2
Mismatch penalty 1 Joining penalty 20
Gap penalty 1.00 Window size 15
Gap size penalty 0.05
Cutoff score 0
Randomization group 0

SEARCH STATISTICS

Scores: Mean Median Standard Deviation
14 15 0.00
Times: CPU 119
00:00:00.00 Total Elapsed
00:00:00.00
Number of residues: 319
Number of sequences searched: 3
Number of scores above cutoff: 3

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Score	Opt.	Sig.	Frame
---------------	-------------	--------	-------	------	------	-------

1. US-09-142-524D-1 Sequence 1, Application U	80	14	14	0.00	0
2. US-09-142-524D-2 Sequence 2, Application U	105	14	14	0.00	0
3. US-09-142-524D-3 Sequence 3, Application U	134	14	14	0.00	0

1. US-09-142-524D-120 (1-15)
US-09-142-524D-1 Sequence 1, Application US/09142524D

Sequence 1, Application US/09142524D
GENERAL INFORMATION:

APPLICANT: Sone, Toshio
APPLICANT: Kume, Akinori
APPLICANT: Dairiki, Kazuo
APPLICANT: Iwama, Akiko
APPLICANT: Kino, Kohsuke
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Diseases
FILE REFERENCE: SPO-103
CURRENT APPLICATION NUMBER: US/09/142,524D
CURRENT FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: PCT/JP97/00740
PRIOR FILING DATE: 1997-03-10
NUMBER OF SEQ ID NOS: 174
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 80
TYPE: PRT
ORGANISM: Cryptomeria japonica

Initial Score = 14 Optimized Score = 14 Significance = 0.00
Residue Identity = 93% Matches = 14 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

MKTVAPOGPNRRVFIKRVSNVLIHGRIDIFASKNFHLOKNTIGTRRISLTKTSGKIASRRVDGIITA
10 20 30 40 50 60 70
X 10 X
GIDIFASKNFHLOKNTIGTRRISLTKTSGKIASRRVDGIITA

YONPASKW
80

2. US-09-142-524D-120 (1-15)
US-09-142-524D-2 Sequence 2, Application US/09142524D

Sequence 2, Application US/09142524D
GENERAL INFORMATION:

APPLICANT: Sone, Toshio
APPLICANT: Kume, Akinori
APPLICANT: Dairiki, Kazuo
APPLICANT: Iwama, Akiko
APPLICANT: Kino, Kohsuke
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Diseases
FILE REFERENCE: SPO-103
CURRENT APPLICATION NUMBER: US/09/142,524D
CURRENT FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: PCT/JP97/00740
PRIOR FILING DATE: 1997-03-10
NUMBER OF SEQ ID NOS: 174
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 105
TYPE: PRT
ORGANISM: Cryptomeria japonica

Initial Score = 14 Optimized Score = 14 Significance = 0.00
Residue Identity = 93% Matches = 14 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X 10 X

GIDIFASKNFHLQKN
 10 20 30 40 X 50 60 70
 MKVTVAENQFGPNRRVFIRKVSNNVTHGRIDIFASKNFHLQKNITGTRRWKNNRIWLOFALITGFTLMGR
 80 90 100
 RLKMWNTIAGYKTFDGRVRVDGIIAAYONPASWK

3. US-09-142-524D-120 (1-15)
 US-09-142-524D-3 Sequence 3, Application US/09142524D

Sequence 3, Application US/09142524D
 GENERAL INFORMATION:

APPLICANT: Some, Toshio
 APPLICANT: Kume, Akinori
 APPLICANT: Dairiki, Kazuo
 APPLICANT: Iwama, Akiko
 APPLICANT: Kino, Kohsuke
 TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease
 FILE REFERENCE: SPO-103

CURRENT APPLICATION NUMBER: US/09/142.524D

CURRENT FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: PCT/JP97/00740

PRIOR FILING DATE: 1997-03-10

NUMBER OF SEQ ID NOS: 174

SOFTWARE: PatentIn version 3.1

SEQ ID NO 3

LENGTH: 134

TYPE: PRT

ORGANISM: Cryptomeria japonica

Initial Score = 14 Optimized Score = 14 Significance = 0.00
 Residue Identity = 93% Matches = 14 Mismatches = 1
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 GIDIFASKNFHLQKN
 10 20 30 40 X 50 60 70
 MKVTVAENQFGPNRRVFIRKVSNNVTHGRIDIFASKNFHLQKNITGTRRWKNNRIWLOFALITGFTLMGR
 80 90 100 110 120 130
 RPLMTIFSGNNINIKLKMENYIAGYKTFDGRVRVDGIIAAYONPASWK

IntelliGenetics

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-09-142-524d-121.res made by jdelaval on Wed 23 Apr 103 17:49:11 PDT

```
Query sequence being compared:US-09-142-524D-121 (1-15)
Number of sequences searched: 3
Number of scores above cutoff: 3
```

Results of the initial comparison of US-09-142-524D-121 (1-15) with:
File : 09142524.pep

Letter	Score
N	1
U	1
M	1
B	1
E	1
R	1
O	1
F	1
S	1
R	1
E	1
C	1
E	1
N	1
C	1
E	1

PARAMETERS

	Unitary	K-truple toining penalty Window size
Similarity matrix	1	20
Mismatch penalty	1.00	2
Gap penalty	0.05	15
Gap size penalty	0	
Cutoff score	0	
Randomization group		

SEARCH STATISTICS

	Mean	Median	Standard Deviation
Scores:	15	16	0.00

```

Times:      CPU      Total Elapsed
00:00:00.00 00:00:00.00

```

```

Number of residues: 319
Number of sequences searched: 3
Number of scores above cutoff: 3

```

The scores below are sorted by initial score. Significance is calculated based on initial score

3 100% similar sequences to the query sequence were found:

Sequence Name	Description	Init.	Opt.	Length	Score	Score	Sig.	Frame
---------------	-------------	-------	------	--------	-------	-------	------	-------

	Sequence 3	Application U	Sequence 2	Application U	Sequence 1	Application U
1. US-09-142-5240-3	134	15	15	0.00	0	0
2. US-09-142-5240-2	105	15	15	0.00	0	0
3. US-09-142-5240-1	80	15	15	0.00	0	0

1. US-09-142-524D-121 (1-15)
US-09-142-524D-3 Sequence 3, Application US/09142524D

Sequence 3, Application US/09142524

APPLICANT: Sone, Toshio
APPLICANT: Kume, Akinori
APPLICANT: Daijiki, Kazuo

APPLICANT: Daitiki, Kazuo
APPLICANT: Iwama, Akiko
APPLICANT: Kino, Kohsuke

TITLE OF INVENTION: PEP
FILE REFERENCE: SPO-103

CURRENT APPLICATION NUMBER: US/03/142,241
CURRENT FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: PCT/JP97/00740

PRIOR FILING DATE: 1997-03-10
NUMBER OF SEQ ID NOS: 174

```
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 134
```

TYPE: PRT
ORGANISM: *Cryptomeria japonica*

```
Initial Score = 15 Optimized Score = 15 Significance = 0.00
Residue Identity = 100% Matches = 15 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0
```

[illegible]

2. US-09-142-524D-121 (1-15)
US-09-142-524D-2 Sequence 2, Application US/09142524D

Sequence 2, Application US/09142524D
GENERAL INFORMATION:

APPLICANT: Sone, Toshio
APPLICANT: Kume, Akinori

APPLICANT: Dairiki, Kazuo
APPLICANT: Iwama, Akiko

APPLICANT: KINO, KONSUKE
TITLE OF INVENTION: Peptide
REFERENCE: SPO-103

FILE REFERENCE: 210 100
CURRENT APPLICATION NUMBER: US/09/142,524D
CURRENT FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: PCT/JP97/00740
PRIOR FILING DATE: 1997-03-10

NUMBER OF SEQ ID NOS: 174
SOFTWARE: PatentIn version 3.1

```
SEQ ID NO 2
LENGTH: 105
TYPE: PPT
```

ORGANISM: *Cryptomeria japonica*

Initial Score	=	15	Optimized Score	=	15	Significance	=	0.0000
Residue Identity	=	100%	Matches	=	15	Mismatches	=	0
Gaps	=	0	Conservative Substitutions	=	0			

X 10 X
ASKNFHLQKNTIGTG
|||||

MKVTVAFNQFGPNRRVFIKRVSNVLIHGRRIDIFASKNFHLQKNTIGTGRMWNRRWLOFAKLTGFTLMGR
10 20 30 40 50 60 70
RLKWPMTYAGYKTFDGRVDGIIIAAYQNPASWK
80 90 100

3. US-09-142-524D-121 (1-15)
US-09-142-524D-1 Sequence 1, Application US/09142524D

Sequence 1, Application US/09142524D
GENERAL INFORMATION:

APPLICANT: Some, Toshio

APPLICANT: Kume, Akinori

APPLICANT: Dairiki, Kazuo

APPLICANT: Iwama, Akiko

APPLICANT: Kino, Kohsuke

TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease

FILE REFERENCE: SPO-103

CURRENT APPLICATION NUMBER: US/09/142,524D

PRIOR FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: PCT/JP97/00740

NUMBER OF SEQ ID NOS: 174

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 80

TYPE: PRT

ORGANISM: Cryptosporidia japonica

Initial Score = 15 Optimized Score = 15 Significance = 0.00
Residue Identity = 100% Matches = 15 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

X 10 X
ASKNPHLOKNTTGTG
10 20 30 40 50 60 70
MKVTVAFNQFGPNRRVFIKRVSNVLIHGRRIDIFASKNFHLQKNTIGTGRMWNRRWLOFAKLTGFTLMGR
YNDPASWK
80

> O <
O | | O < Intelligenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-09-142-524d-131.res made by jdelaval on Wed 23 Apr 103 17:49:28-PDT.

Query sequence being compared: US-09-142-524D-131 (1-15)
Number of sequences searched: 3
Number of scores above cutoff: 3

Results of the initial comparison of US-09-142-524D-131 (1-15) with:
File : 09142524.pep

```

100-
N -
U - 50-
M -
B -
E -
R -
O -
F - 10-
S -
S -
Q -
U -
E -
N -
C -
E -
S -
SCORE 0 1 2 3 5 6 8 9 11 12 14
STDEV 0 0 0 0 0 0 0 0 0 0 0 0

```

PARAMETERS

Similarity matrix Unitary K-tuple 2
Mismatch penalty 1 Joining penalty 20
Gap penalty 1.00 Window size 15
Gap size penalty 0.05
Cutoff score 0
Randomization group 0

SEARCH STATISTICS

Scores: Mean 6 Median 3 Standard Deviation 6.93
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00

Number of residues: 319
Number of sequences searched: 3
Number of scores above cutoff: 3

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Score	Opt. Frame
1. US-09-142-524D-3	Sequence 3, Application U	134	14	1.15 0
2. US-09-142-524D-1	Sequence 1, Application U	80	2	4 -0.58 0
3. US-09-142-524D-2	Sequence 2, Application U	105	2	4 -0.58 0

1. US-09-142-524D-131 (1-15)
US-09-142-524D-3 Sequence 3, Application US/09142524D

Sequence 3, Application US/09142524D
GENERAL INFORMATION:
APPLICANT: Sone, Toshio
APPLICANT: Kume, Akimori
APPLICANT: Dairiki, Kazuo
APPLICANT: Iwama, Akiko

TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease
FILE REFERENCE: SPO-103
CURRENT APPLICATION NUMBER: US/09/142,524D
CURRENT FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: PCT/JP97/00740
PRIOR FILING DATE: 1997-03-10

NUMBER OF SEQ ID NOS: 174
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 134

TYPE: PRT
ORGANISM: Cryptomeria japonica

Initial Score = 14 Optimized Score = 14 Significance = 1.15
Residue Identity = 93% Matches = 14 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

MKVTVAFNPGFNRVFIKRVSNVTHGRIDIPASKNFHLQKNKTGTGRWKNRIMQLQFALKTGPTLMGR
10 20 30 40 50 60 70

X
SRAEVSYYHVGAKF
RPLWTFSGNNMTKLKPMYTAGYKTFDGRRAEVSYYHVGAKFIRRVDDGIIAAYONPASWK
80 90 100 X 110 X 120 130

2. US-09-142-524D-131 (1-15)
US-09-142-524D-1 Sequence 1, Application US/09142524D

Sequence 1, Application US/09142524D
GENERAL INFORMATION:
APPLICANT: Sone, Toshio
APPLICANT: Kume, Akimori
APPLICANT: Dairiki, Kazuo
APPLICANT: Iwama, Akiko

TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease
FILE REFERENCE: SPO-103
CURRENT APPLICATION NUMBER: US/09/142,524D
CURRENT FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: PCT/JP97/00740
PRIOR FILING DATE: 1997-03-10

NUMBER OF SEQ ID NOS: 174
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 80

TYPE: PRT
ORGANISM: Cryptomeria japonica

Initial Score = 2 Optimized Score = 4 Significance = -0.58
Residue Identity = 26% Matches = 4 Mismatches = 11
Gaps = 0 Conservative Substitutions = 0

X 10 X
SRAEVSYPHNGAKF
10 X 20 30 40 50 60 70
MKVTVAFNQFGPNRRVETIKRVSNVLIHGRRIDIFASKNPHLQKNTIGTRRWKNNRIWLOFAKLTGFTLMGR

YONPASWK
80

3. US-09-142-524D-131 (1-15)
US-09-142-524D-2 Sequence 2, Application US/09142524D

Sequence 2, Application US/09142524D
GENERAL INFORMATION:

APPLICANT: Sone, Toshio
APPLICANT: Kume, Akinori
APPLICANT: Dairiki, Kazuo
APPLICANT: Iwama, Aiko
APPLICANT: Kino, Kohsuke
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease
FILE REFERENCE: SPO-103
CURRENT APPLICATION NUMBER: US/09/142,524D
CURRENT FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: PCT/JP97/00740
PRIOR FILING DATE: 1997-03-10
NUMBER OF SEQ ID NOS: 174
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 105
TYPE: PRT
ORGANISM: Cryptomeria japonica

Initial Score = 2 Optimized Score = 4 Significance = -0.58
Residue Identity = 26% Matches = 4 Mismatches = 11
Gaps = 0 Conservative Substitutions = 0

X 10 X
SRAEVSYPHNGAKF
10 X 20 30 40 50 60 70
MKVTVAFNQFGPNRRVETIKRVSNVLIHGRRIDIFASKNPHLQKNTIGTRRWKNNRIWLOFAKLTGFTLMGR
RLKMPMTIAGYKTFDGRVDGIIAAYQNPASWK
80 90 100

IntelliGenetics

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

results file us-09-142-524d-152.res made by jdelaval on Wed 23 Apr 103 17:49:45-PDT.

Query sequence being compared:	US-09-142-524D-152 (1-15)
Number of sequences searched:	3
Number of scores above cutoff:	3

Results of the initial comparison of US-09-142-524D-152 (1-15) with
File : 09142524.pep

[illegible]

PARAMETERS

Unitary	2
K-tuple	20
Joining penalty	20
Window size	15
Mismatch penalty	1
Gap penalty	1.00
Gap size penalty	0.05
Cutoff score	0
Randomization group	0

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	6	4	5.20
Times:	CPU	Total Elapsed	
	00:00:00.00	00:00:00.00	

Number of residues:	319
Number of sequences searched:	3
Number of scores above cutoff:	3

The scores below are sorted by initial score. Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found

The list of best scores is:

Sequence Name	Description	Init. Opt.	Length Score	score	sig.	Frames
---------------	-------------	------------	--------------	-------	------	--------

	1	2	3	4
1. US-09-142-524D-1 Sequence 1, Application U	mean	12	1.15	0
2. US-09-142-524D-2 Sequence 2, Application U	standard deviation	80	12	0
3. US-09-142-524D-3 Sequence 3, Application U	from mean	105	4	-0.58
		134	3	-0.58

1. US-09-142-524D-152 (1-15)
US-09-142-524D-1 Sequence 1, Application US/09142524E

Sequence 1, Application US/09142524D
GENERAL INFORMATION:

APPLICANT: Some, Toshio
 APPLICANT: Kume, Akinori
 APPLICANT: Dairiki, Kazuo
 APPLICANT: Iwama, Akiro
 APPLICANT: Kino, Kohnsuke
 TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease
 FILE REFERENCE: S50-103

CURRENT APPLICATION NUMBER: US-09/142,524D
 CURRENT FILING DATE: 1998-09-09
 PRIOR APPLICATION NUMBER: PCT/JP97/00740
 PRIOR FILING DATE: 1997-03-10
 NUMBER OF SEQ ID NOS: 174
 *SOFTWARE: PatentIn version 3.1
 SEQ ID NO: 1

TYPE: PRT
ORGANISM: *Cryptomeria japonica*

Initial Score	=	12	Optimized Score	=	12	Significance	=	1.15
Residue Identity	=	80%	Matches	=	12	Mismatches	=	3
Gaps	=	0	Conservative Substitutions	=	0			

MKVTVANQGNRRRVFIKRVSNVTHGRRIIDFASKNHLQNTTGTCTRTSLKTSKISRSRVGGIIAA
 10 20 30 40 50 60 70

YQNPAWK
80

2. US-09-142-524D-152 (1-15)
US-09-142-524D-2 Sequence 2, Application US/09142524D

Sequence 2, Application US/09142524D
GENERAL INFORMATION:

APPLICANT: Some, Toshio
 APPLICANT: Kume, Akinori
 APPLICANT: Dairiki, Kazuo
 APPLICANT: Iwama, Akiko
 APPLICANT: Kino, Koshuke
 TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Diseases
 FILE REFERENCE: SPO-103
 CURRENT APPLICATION NUMBER: US/09/142,524D
 PRIOR FILING DATE: 1998-09-09
 PRIOR APPLICATION NUMBER: PCT/jp97/00740
 PRIOR FILING DATE: 1997-03-10
 NUMBER OF SEQ ID NOS: 174
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 2
 LENGTH: 105

ORGANISM: *Cryptomeria japonica*

Initial Score	=	3	Optimized Score	=	4	Significance	=	-0.56
Residue Identity	=	27%	Matches	=	5	Mismatches	=	10
Gaps	=	3	Conservative Substitutions	=	0			

X
 10 X
 LSDISL---KLTSGKIAS
 10 20 30 40 50 60 70
 MKTVAFNPGFGRNRVFIKRVSNVTHGRRIDIFASKNFHLOKNTIGTRGRKNNRIWLOFAKLTGFTLMGR

RLKMPMYIAGYKTFDGRRVDCIIAAYQNPASWK
 80 90 100

3. US-09-142-524D-152 (1-15)
 US-09-142-524D-3 Sequence 3, Application US/09142524D

Sequence 3, Application US/09142524D

GENERAL INFORMATION:
 APPLICANT: Sone, Toshio
 APPLICANT: Kume, Akimori
 APPLICANT: Dairiki, Kazuo
 APPLICANT: Iwama, Akiko
 APPLICANT: Kuno, Kohsuke
 TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease
 FILE REFERENCE: SPO-103
 CURRENT APPLICATION NUMBER: US/09/142,524D
 CURRENT FILING DATE: 1998-09-09
 PRIOR APPLICATION NUMBER: PCT/JP97/00740
 PRIOR FILING DATE: 1997-03-10
 NUMBER OF SEQ ID NOS: 174
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 3
 LENGTH: 134
 TYPE: PRT
 ORGANISM: Cryptosporidia japonica

Initial Score = 3 Optimized Score = 4 Significance = -0.58
 Residue Identity = 27% Matches = 5 Mismatches = 10
 Gaps = 3 Conservative Substitutions = 0

X
 10 X
 LSDISL---KLTSGKIAS
 10 20 30 40 50 60 70
 MKTVAFNPGFGRNRVFIKRVSNVTHGRRIDIFASKNFHLOKNTIGTRGRKNNRIWLOFAKLTGFTLMGR
 80 90 100 110 120 130
 RPLMIIFSGNMNIKLMKMPYIAGYKTFDGRRAEVSIVHNGAKFIRVDGIIAAYQNPASWK

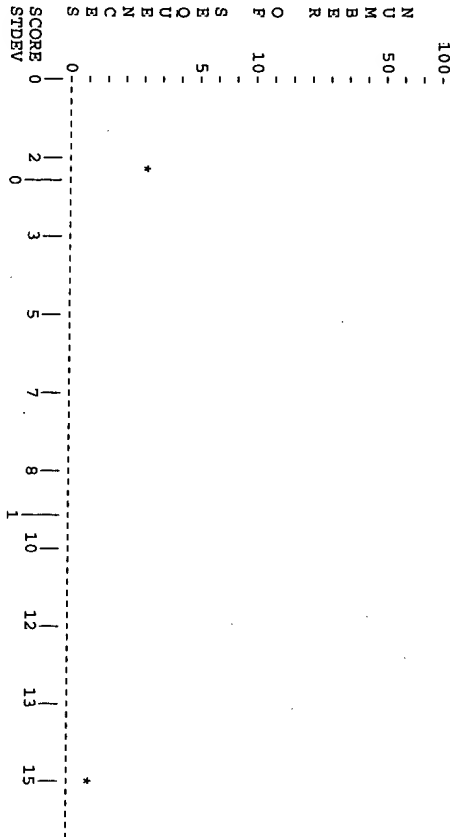
> O <
O | 10 Intelligenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-09-142-524d-28.res made by jdelaval on Wed 23 Apr 103 17:45:42-PDT.

Query sequence being compared: US-09-142-524D-28 (1-15)
Number of sequences searched: 3
Number of scores above cutoff: 3

Results of the initial comparison of US-09-142-524D-28 (1-15) with:
File : 09142524.pep



PARAMETERS

Similarity matrix Unitary 1 K-tuple 2
Mismatch penalty 1 Joining penalty 20
Gap penalty 1.00 Window size 15
Gap size penalty 0.05
Cutoff score 0
Randomization group 0

SEARCH STATISTICS

Scores: Mean Median Standard Deviation
6 3- 7.51
Times: CPU Total Elapsed
00:00:00.00 00:00:00.00
Number of residues: 319
Number of sequences searched: 3
Number of scores above cutoff: 3

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% similar sequence to the query sequence was found:

Sequence Name	Description	Init. Opt.	Sig. Frame
		Length Score	Score

1. US-09-142-524D-3 Sequence 3, Application U 134 15 15 1.20 0

The list of other best scores is:

Sequence Name	Description	Length	Score	Opt.	Sig.	Frame
2. US-09-142-524D-1	Sequence 1, Application U	80	2	3	-0.53	0
3. US-09-142-524D-2	Sequence 2, Application U	105	2	3	-0.53	0

1. US-09-142-524D-28 (1-15)
US-09-142-524D-3 Sequence 3, Application US/09142524D

Sequence 3, Application US/09142524D
GENERAL INFORMATION:

APPLICANT: Sone, Toshio
APPLICANT: Kume, Akimori
APPLICANT: Dairiki, Kazuo
APPLICANT: Iwama, Akiko
APPLICANT: Kino, Kohsuke
TITLE OF INVENTION: Peptide-Based Immunoth.
FILE REFERENCE: SPO-103
CURRENT APPLICATION NUMBER: US/09/142,524D
CURRENT FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: PCT/JP97/00740
PRIOR FILING DATE: 1997-03-10
NUMBER OF SEQ ID NOS: 174
SOFTWARE: Patent version 3.1
SEQ ID NO 3
LENGTH: 134
TYPE: PRT
ORGANISM: Cryptomeria japonica

Initial Score = 15 Optimized Score = 15 Significance = 1.20
Residue Identity = 100% Matches = 15 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

MKTVAAGGPNRRVRIKRVSVITIGRRIDIFASKNFHLQKNTIGTGRMKNNRIMLQFAKLTGFTLMGR
10 20 30 40 50 60 70
X 10 X
RPLWITFGNNITKL
RPLWITFGNNITKLKPMYTAGYTFDGRRAEVSIVHNGAKFIRVDGIIAAYONPASWK
X 80 X 90 100 110 120 130

2. US-09-142-524D-28 (1-15)
US-09-142-524D-1 Sequence 1, Application US/09142524D

Sequence 1, Application US/09142524D
GENERAL INFORMATION:

APPLICANT: Sone, Toshio
APPLICANT: Kume, Akimori
APPLICANT: Dairiki, Kazuo
APPLICANT: Iwama, Akiko
APPLICANT: Kino, Kohsuke
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Diseases
FILE REFERENCE: SPO-103
CURRENT APPLICATION NUMBER: US/09/142,524D
CURRENT FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: PCT/JP97/00740
PRIOR FILING DATE: 1997-03-10
NUMBER OF SEQ ID NOS: 174
SOFTWARE: Patent version 3.1
SEQ ID NO 1
LENGTH: 80
TYPE: PRT
ORGANISM: Cryptomeria japonica

Initial Score = 2 Optimized Score = 3 Significance = -0.53

Residue Identity = 23% Matches = 4 Mismatches = 11
Gaps = 2 Conservative Substitutions = 0

X
RPLWIFSGN--NNIKL
X
MKTVAENQGPNNRRVFIKRVSNVTHGRRIDIPASKNPHLQKNTIGTGRRLSKLTSGKIASRRVDGIILAA
X 10 20 30 40 50 60 70
YONPASWK
80

3. US-09-142-524D-28 (1-15)
US-09-142-524D-2 Sequence 2, Application US/09142524D

Sequence 2, Application US/09142524D
GENERAL INFORMATION:

APPLICANT: Sone, Toshio
APPLICANT: Kume, Akio
APPLICANT: Dairiki, Kazuo
APPLICANT: Iwama, Akiko
APPLICANT: Kino, Kohsuke
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease
FILE REFERENCE: SPQ-103
CURRENT APPLICATION NUMBER: US/09/142,524D
CURRENT FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: PCT/JP97/00740
PRIOR FILING DATE: 1997-03-10
NUMBER OF SEQ ID NOS: 174
SOFTWARE: Patent version 3.1
SEQ ID NO 2

LENGTH: 105
TYPE: PRT

ORGANISM: Cryptomeria japonica

Initial Score = 2 Optimized Score = 3 Significance = -0.53
Residue Identity = 23% Matches = 4 Mismatches = 11
Gaps = 2 Conservative Substitutions = 0

X
RPLWIFSGN--NNIKL
X
MKTVAENQGPNNRRVFIKRVSNVTHGRRIDIPASKNPHLQKNTIGTGRRLSKLTSGKIASRRVDGIILAA
X 10 20 30 40 50 60 70
RLKMPYIAGYKTFDGRRRVDGIILAA YONPASWK
80 90 100

> 0
0/10 Intelligenetics
> 0 <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-09-142-524d-29.res made by jdelaval on Wed 23 Apr 103 17:46:13-PDT.

Query sequence being compared: US-09-142-524D-29 (1-15)
Number of sequences searched: 3
Number of scores above cutoff: 3

Results of the initial comparison of US-09-142-524D-29 (1-15) with:
File : 09142524.pep

100-
N -
U -
M -
B -
E -
R -
O -
F -
S -
E -
C -
S -
SCORE 0 2 3 5 7 8 10 12 13 15
STDEV 0 1 1 1 1 1 1 1 1 1

PARAMETERS

Similarity matrix Unitary 1 K-tuple 2
Mismatch penalty 1 Joining penalty 20
Gap penalty 1.00 Window size 15
Gap size penalty 0.05
Cutoff score 0
Randomization group 0

SEARCH STATISTICS

Scores: Mean 7 Median 3 Standard Deviation 6.66
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00

Number of residues: 319
Number of sequences searched: 3
Number of scores above cutoff: 3

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% similar sequence to the query sequence was found:

Sequence Name	Description	Init. Opt.	Length	Score	Sig. Frame
---------------	-------------	------------	--------	-------	------------

1. US-09-142-524D-3 Sequence 3, Application U 134 15 15 1.20 0

The list of other best scores is:

Sequence Name	Description	Length	Score	Sig. Frame
2. US-09-142-524D-2 Sequence 2, Application U		105	6	7 -0.15 0
3. US-09-142-524D-1 Sequence 1, Application U		80	2	3 -0.75 0

1. US-09-142-524D-29 (1-15)

US-09-142-524D-3 Sequence 3, Application US/09142524D

Sequence 3, Application US/09142524D
GENERAL INFORMATION:

APPLICANT: Some, Toshio

APPLICANT: Kume, Akinori

APPLICANT: Dairiki, Kazuo

APPLICANT: Iwama, Akiko

APPLICANT: Kino, Kohsuke

TITLE OF INVENTION: Peptide-based Immunotherapeutic Agent for Treating Allergic Disease

FILE REFERENCE: SPO-103

CURRENT APPLICATION NUMBER: US/09/142,524D

CURRENT FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: PCT/JP97/00740

PRIOR FILING DATE: 1997-03-10

NUMBER OF SEQ ID NOS: 174

SOFTWARE: Patent in version 3.1

SEQ ID NO 3

LENGTH: 134

TYPE: PRT

ORGANISM: Cryptosporidia japonica

Initial Score = 15 Optimized Score = 15 Significance = 1.20
Residue Identity = 100% Matches = 15 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

MKVTVAENQFGPNRRVFIKRVSNVLIHGRRIDIPASKNPHLOKNTIGTGRMKNNRIWLOFAKLTGFTLMGR
10 20 30 40 50 60 70

X 10 X
IFSGNNNITLKKPMY
|||||

RPLWTFSGNNNITLKKPMYIAGYKTFDGRRAEVSYYVWGAKFIRRDGIIAAYONPASWK
80 90 X 100 110 120 130

2. US-09-142-524D-29 (1-15)

US-09-142-524D-2 Sequence 2, Application US/09142524D

Sequence 2, Application US/09142524D
GENERAL INFORMATION:

APPLICANT: Some, Toshio

APPLICANT: Kume, Akinori

APPLICANT: Dairiki, Kazuo

APPLICANT: Iwama, Akiko

APPLICANT: Kino, Kohsuke

TITLE OF INVENTION: Peptide-based Immunotherapeutic Agent for Treating Allergic Disease

FILE REFERENCE: SPO-103

CURRENT APPLICATION NUMBER: US/09/142,524D

CURRENT FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: PCT/JP97/00740

PRIOR FILING DATE: 1997-03-10

NUMBER OF SEQ ID NOS: 174

SOFTWARE: Patent in version 3.1

SEQ ID NO 2

LENGTH: 105

TYPE: PRT

ORGANISM: Cryptosporidia japonica

Initial Score = 6 Optimized Score = 7 Significance = -0.15

Residue Identity = 46% Matches = 7 Mismatches = 8
 Gaps = 0 Conservative Substitutions = 0

MKVTVAFNQFGPNRRVFIKRVSNVTHGRRIDIPASKNFHLQKNITIGTGRMRKNNRIMLQFAKLTGFTLMGR
 10 20 30 40 50 60 70
 X
 IFSGNMNI

10 X
 KLTKMPMY
 |||||
 RLKMPMYIAGYKTFEDGRRVDGIITAYGNPMSWK
 80 90 100

3. US-09-142-524D-29 (1-15)
 US-09-142-524D-1 Sequence 1, Application US/09142524D

Sequence 1, Application US/09142524D
 GENERAL INFORMATION:

APPLICANT: Sone, Toshio

APPLICANT: Kume, Akimori

APPLICANT: Dairiki, Kazuo

APPLICANT: Iwama, Akiko

APPLICANT: Kino, Kohsuke

TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease

FILE REFERENCE: SPO-103

CURRENT APPLICATION NUMBER: US/09/142,524D

CURRENT FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: PCT/JP97/00740-

PRIOR FILING DATE: 1997-03-10

NUMBER OF SEQ ID NOS: 174

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 80

TYPE: PRT

ORGANISM: Cryptosporidia japonica

Initial Score = 2 Optimized Score = 3 Significance = -0.75
 Residue Identity = 20% Matches = 3 Mismatches = 12
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 IFSGNMNITKLKMPMY
 |
 MKVTVAFNQFGPNRRVFIKRVSNVTHGRRIDIPASKNFHLQKNITIGTGRRIISLKLTSKIASRRVDGIITAA
 10 20 X 30 40 50 60 70
 YNDPMSWK
 80

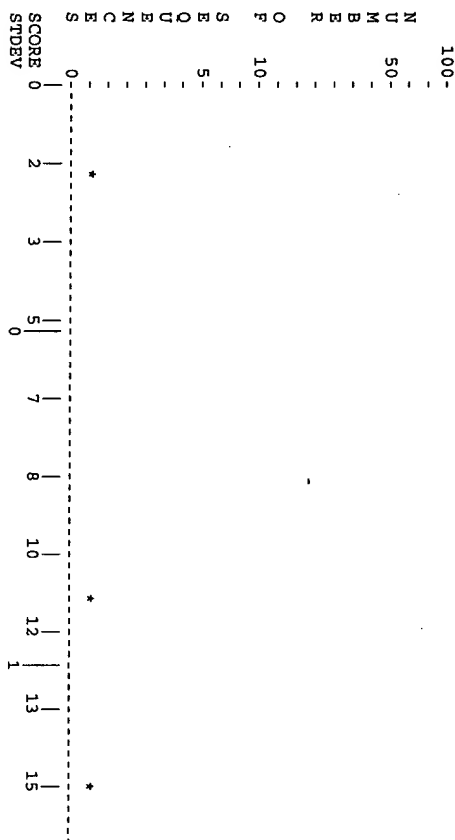
> O
O I 0 IntelliGenetics
> O <

PasteDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-09-142-524d-30.res made by jdelaval on Wed 23 Apr 103 17:46:35-PDT.

Query sequence being compared: US-09-142-524D-30 (1-15)
Number of sequences searched: 3
Number of scores above cutoff: 3

Results of the initial comparison of US-09-142-524D-30 (1-15) with:
File : 09142524.pep



SEARCH STATISTICS

Scores: Mean 9 Median 3 Standard Deviation 6.66
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00
Number of residues: 319
Number of sequences searched: 3
Number of scores above cutoff: 3

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% similar sequence to the query sequence was found:

Sequence Name	Description	Int. Opt.	Length	Score	Sig. Frame
---------------	-------------	-----------	--------	-------	------------

1. US-09-142-524D-3 Sequence 3, Application U 134 15 15 0.90 0
The list of other best scores is:

Sequence Name	Description	Length	Score	Opt.	Sig. Frame
2. US-09-142-524D-2 Sequence 2, Application U	105 11 12 0.30 0				
3. US-09-142-524D-1 Sequence 1, Application U	80 2 4 -1.05 0				

1. US-09-142-524D-30 (1-15)
US-09-142-524D-3 Sequence 3, Application US/09142524D

Sequence 3, Application US/09142524D
GENERAL INFORMATION:

APPLICANT: Sone, Toshio
APPLICANT: Kume, Akinori
APPLICANT: Dairiki, Kazuo
APPLICANT: Iwama, Akiyo
APPLICANT: Kino, Kohsuke
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease
FILE REFERENCE: SPO-103
CURRENT APPLICATION NUMBER: US/09/142,524D
CURRENT FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: PCT/J97/00740
PRIOR FILING DATE: 1997-03-10
NUMBER OF SEQ ID NOS: 174
SOFTWARE: Patent in version 3.1
SEQ ID NO 3
LENGTH: 134
TYPE: PRT
ORGANISM: Cryptomeria japonica

Initial Score = 15 Optimized Score = 15 Significance = 0.90
Residue Identity = 100% Matches = 15 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

MKVTVAENQSGPNRRVPIKRVSNVLIHGRIDIFASXNPHLQKNTIGTGRWKNRIMLQFAPLGTGFTLMGR
10 20 30 40 50 60 70
X 10 X
MNIKIKMPMYIAGYK
RPLWTIFSGNNIKLKMPLYAGYKTFDGRRAEVSYYHVGAKFIRVDGIITAYONPASWK
80 X 90 100 110 120 130

2. US-09-142-524D-30 (1-15)
US-09-142-524D-2 Sequence 2, Application US/09142524D

Sequence 2, Application US/09142524D
GENERAL INFORMATION:

APPLICANT: Sone, Toshio
APPLICANT: Kume, Akinori
APPLICANT: Dairiki, Kazuo
APPLICANT: Iwama, Akiyo
APPLICANT: Kino, Kohsuke
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease
FILE REFERENCE: SPO-103
CURRENT APPLICATION NUMBER: US/09/142,524D
CURRENT FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: PCT/J97/00740
PRIOR FILING DATE: 1997-03-10
NUMBER OF SEQ ID NOS: 174
SOFTWARE: Patent in version 3.1
SEQ ID NO 2
LENGTH: 105
TYPE: PRT
ORGANISM: Cryptomeria japonica

> 0 <
0 | 0 IntellGenetics
> 0 <

PaetDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-09-142-524d-31.res made by jdelaval on Wed 23 Apr 103 17:46:55-PDT.

Query sequence being compared: US-09-142-524D-31 (1-15)
Number of sequences searched: 3
Number of scores above cutoff: 3

Results of the initial comparison of US-09-142-524D-31 (1-15) with:
File : 09142524.pep

100-
N -
U 50-
M -
B -
E -
R -
O 10-
S -
E 5-
O -
U -
N -
C -
E -
S -
SCORE 0 2 3 5 7 8 10 12 13 15
STDEV

PARAMETERS

Similarity matrix Unitary 1 K-tuple 2
Mismatch penalty 1.00 Joining penalty 20
Gap size penalty 0.05 Window size 15
Cutoff score 0
Randomization group 0

SEARCH STATISTICS

Scores: Mean Median Standard Deviation
10 3 7.51
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00
Number of residues: 319
Number of sequences searched: 3
Number of scores above cutoff: 3

The scores below are sorted by initial score.
Significance is calculated based on initial score.

2 100% similar sequences to the query sequence were found:

Sequence Name	Description	Int. Opt.	Length	Score	Sig. Frame
---------------	-------------	-----------	--------	-------	------------

1. US-09-142-524D-3 Sequence 3, Application U 134 15 15 0.67 0
2. US-09-142-524D-2 Sequence 2, Application U 105 15 15 0.67 0
The list of other best scores is:

Sequence Name	Description	Length	Score	Int. Opt.	Sig. Frame
---------------	-------------	--------	-------	-----------	------------

3. US-09-142-524D-1 Sequence 1, Application U 80 2 3 -1.07 0
**** 1 standard deviation below mean ****

1. US-09-142-524D-31 (1-15)
US-09-142-524D-3 Sequence 3, Application US/09142524D

Sequence 3, Application US/09142524D
GENERAL INFORMATION:

APPLICANT: Sone, Toshio
APPLICANT: Kume, Akimori
APPLICANT: Dairiki, Kazuo
APPLICANT: Iwama, Akiko
APPLICANT: Kino, Kohsuke
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease
FILE REFERENCE: SPO-103
CURRENT APPLICATION NUMBER: US/09/142,524D
CURRENT FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: PCT/JP97/00740
PRIOR FILING DATE: 1997-03-10
NUMBER OF SEQ ID NOS: 174
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 134
TYPE: PRT
ORGANISM: Cryptomeria japonica

Initial Score = 15 Optimized Score = 15 Significance = 0.67
Residue Identity = 100% Matches = 15 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

MKVTVAFNQGPNRRVFIKRVSNVLIHGRRIDIFASKNPHLOKNTIGTGRWKNRIMLOPAKLTGFTLMGR
10 20 30 40 50 60 70

X 10 X
KMPMYAGYKTFDGR
RPLWTFSGNNMNIKLKMPMYIAGYKTFDGRRAEVSYYVANGAKFIRRDGIIAAYONPASWK
80 90 100 X 110 120 130

2. US-09-142-524D-31 (1-15)
US-09-142-524D-2 Sequence 2, Application US/09142524D

Sequence 2, Application US/09142524D
GENERAL INFORMATION:

APPLICANT: Sone, Toshio
APPLICANT: Kume, Akimori
APPLICANT: Dairiki, Kazuo
APPLICANT: Iwama, Akiko
APPLICANT: Kino, Kohsuke
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease
FILE REFERENCE: SPO-103
CURRENT APPLICATION NUMBER: US/09/142,524D
CURRENT FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: PCT/JP97/00740
PRIOR FILING DATE: 1997-03-10
NUMBER OF SEQ ID NOS: 174
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 105
TYPE: PRT
ORGANISM: Cryptomeria japonica

Initial Score = 15 Optimized Score = 15 Significance = 0.67
 Residue Identity = 100% Matches = 15 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

MKVTVAENQFGPNRRVFVKRVSNIHGRIDIFASKNFHLQKNTIGTGRRWKNNRIWLFALYGFLLMGR
 10 20 30 40 50 60 70

X 10 X
 KMPMYIAGYKTFDGR
 |||||
 RLKMPMYIAGYKTFDGRVDGIITAAVQNPASWK
 X 80 90 100

3. US-09-142-524D-31 (1-15)
 US-09-142-524D-1 Sequence 1, Application US/09142524D

Sequence 1, Application US/09142524D
 GENERAL INFORMATION:

APPLICANT: Sone, Toshio

APPLICANT: Kume, Akinori

APPLICANT: Dairiki, Kazuo

APPLICANT: Iwama, Akiko

APPLICANT: Kino, Kohsuke
 TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease

FILE REFERENCE: SPO-103
 CURRENT APPLICATION NUMBER: US/09/142,524D

PRIOR APPLICATION NUMBER: PCT/JP97/00740

PRIOR FILING DATE: 1997-03-10

NUMBER OF SEQ ID NOS: 174

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1
 LENGTH: 80

TYPE: PRT
 ORGANISM: Cryptosporidia japonica

Initial Score = 2 Optimized Score = 3 Significance = -1.07
 Residue Identity = 20% Matches = 4 Mismatches = 11
 Gaps = 5 Conservative Substitutions = 0

X 10 X
 KMPMYIAGYK----TFDGR
 ||
 MKVTVAENQFGPNRRVFVKRVSNIHGRIDIFASKNFHLQKNTIGTGRRIISLKTSGKIASRRVDGIITAA
 10 20 30 40 50 60 70
 YQNPASWK
 80

> 0 <
O | O IntelliGenetics
> 0 <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-09-142-524d-36.res made by jdelaval on Wed 23 Apr 103 17:47:17-PDT.

Query sequence being compared: US-09-142-524D-36 (1-15)
Number of sequences searched: 3
Number of scores above cutoff: 3

Results of the initial comparison of US-09-142-524D-36 (1-15) with:
File : 09142524.pep

```

100-
N   -
U   50-
M   -
B   -
R   -
O   10-
S   -
E   5-
U   -
E   -
N   -
C   -
E   -
S   -
SCORE 0 1 3 4 6 7 9 10 12 13
STDVY

```

PARAMETERS

Similarity matrix Unitary K-tuple
Mismatch penalty 1 Joining penalty 2
Gap penalty 1.00 Window size 20
Gap size penalty 0.05
Cutoff score 0
Randomization group 0

SEARCH STATISTICS

Scores: Mean 13 Median 14 Standard Deviation 0.00
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00

Number of residues: 319
Number of sequences searched: 3
Number of scores above cutoff: 3

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Score	Int. Opt. Score	Sig. Frame
---------------	-------------	--------	-------	-----------------	------------

1. US-09-142-524D-1	Sequence 1, Application U	80	13	13	0.00 0
2. US-09-142-524D-2	Sequence 2, Application U	105	13	13	0.00 0
3. US-09-142-524D-3	Sequence 3, Application U	134	13	13	0.00 0

1. US-09-142-524D-36 (1-15)
US-09-142-524D-1 Sequence 1, Application US/09142524D

Sequence 1, Application US/09142524D
GENERAL INFORMATION:

APPLICANT: Sone, Toshio
APPLICANT: Kume, Akinori
APPLICANT: Dairiki, Kazuo

APPLICANT: Iwama, Akiho

APPLICANT: Kino, Kohsuke

TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease

FILE REFERENCE: SPO-103

CURRENT APPLICATION NUMBER: US/09/142,524D

PRIOR FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: PCT/JP97/00740

PRIOR FILING DATE: 1997-03-10

NUMBER OF SEQ ID NOS: 174

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 80

TYPE: PRT

ORGANISM: Cryptomeria japonica

Initial Score = 13 Optimized Score = 13 Significance = 0.00
Residue Identity = 86% Matches = 13 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

PCVFIRKRVSNVYIHG
MKATVAFNQFGNRRVFIKRVSNVYIHGRIDIPASKNFHLQKNTIGRRRLSLKTSKIASRRVDGIIAA
10 X 20 30 40 50 60 70

YONPASWK
80

2. US-09-142-524D-36 (1-15)
US-09-142-524D-2 Sequence 2, Application US/09142524D

Sequence 2, Application US/09142524D
GENERAL INFORMATION:

APPLICANT: Sone, Toshio

APPLICANT: Kume, Akinori

APPLICANT: Dairiki, Kazuo

APPLICANT: Iwama, Akiho

APPLICANT: Kino, Kohsuke

TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease

FILE REFERENCE: SPO-103

CURRENT APPLICATION NUMBER: US/09/142,524D

CURRENT FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: PCT/JP97/00740

PRIOR FILING DATE: 1997-03-10

NUMBER OF SEQ ID NOS: 174

SOFTWARE: PatentIn version 3.1

SEQ ID NO 2

LENGTH: 105

TYPE: PRT

ORGANISM: Cryptomeria japonica

Initial Score = 13 Optimized Score = 13 Significance = 0.00
Residue Identity = 86% Matches = 13 Mismatches = 2
Gaps = 0 Conservative Substitutions = 0

X 10 X

PCVFIKRVSNVTHG
 |||||
 10 X 20 30 40 50 60 70
 MKVTVAFNQFGPNRVRVFIKRVSNVTHGRIIDIFASKNFHLQKNVTIGTRRMKNRRIWLOFAKLTGFTLMGR
 80 90 100
 RLKMPMYIAGYKTFDGRRRVDGIIAAYQNPASWK

3. US-09-142-524D-36 (1-15)
 US-09-142-524D-3 Sequence 3, Application US/09142524D

Sequence 3, Application US/09142524D

GENERAL INFORMATION:

APPLICANT: Sone, Toshio

APPLICANT: Kume, Akinori

APPLICANT: Daitiki, Kazuo

APPLICANT: Iwama, Akiko

APPLICANT: Kino, Kousuke

TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Diseases

FILE REFERENCE: SPO-103

CURRENT APPLICATION NUMBER: US/09/142,524D

PRIOR FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: PCT/JP97/00740

PRIOR FILING DATE: 1997-03-10

NUMBER OF SEQ ID NOS: 174

SOFTWARE: PatentIn version 3.1

SEQ ID NO 3

LENGTH: 134

TYPE: PRT

ORGANISM: Cryptosporidia japonica

Initial Score = 13

Residue Identity = 86%

Gaps = 0

Optimized Score = 13

Matches = 13

Significance = 0.00

Conservative Substitutions = 2

Mismatches = 0

PCVFIKRVSNVTHG

|||||

10 X 20 30 40 50 60 70

MKVTVAFNQFGPNRVRVFIKRVSNVTHGRIIDIFASKNFHLQKNVTIGTRRMKNRRIWLOFAKLTGFTLMGR

80 90 100 110 120 130

RLKMPMYIAGYKTFDGRRAEVSYYHNGAKFIRRVDDGIIAAYQNPASWK

> O <
 O | 10 IntelliGenetics
 > O <
 FastDB - Fast Pairwise Comparison of Sequences
 Release 5.4

Results file us-09-142-524d-57.res made by jdelaval on Wed 23 Apr 103 17:47:36-PDT.

Query sequence being compared: US-09-142-524D-57 (1-15)
 Number of sequences searched: 3
 Number of scores above cutoff: 3

Results of the initial comparison of US-09-142-524D-57 (1-15) with:
 File: 09142524.pep

100-
 N -
 U - 50-
 M -
 B -
 E -
 R -
 O -
 F 10-
 S -
 E 5-
 U -
 Q -
 N -
 E -
 C -
 S -
 S 0-
 SCORE 0 1 3 4 6 7 9 10 12 13
 STDEV

PARAMETERS

Similarity matrix Unitary 2
 Mismatch penalty 1 Joining penalty 20
 Gap penalty 1.00 Window size 15
 Gap size penalty 0.05
 Cutoff score 0
 Randomization group 0

SEARCH STATISTICS

Scores: Mean 13 Median 14 Standard Deviation 0.00
 Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00
 Number of residues: 319
 Number of sequences searched: 3
 Number of scores above cutoff: 3

The scores below are sorted by initial score.
 Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Score	Int. Opt. Score	Sig. Frame
1. US-09-142-524D-1	Sequence 1, Application U	80	13	13	0.00 0
2. US-09-142-524D-2	Sequence 2, Application U	105	13	13	0.00 0
3. US-09-142-524D-3	Sequence 3, Application U	134	13	13	0.00 0

1. US-09-142-524D-57 (1-15)
 US-09-142-524D-1 Sequence 1, Application US/09142524D

Sequence 1, Application US/09142524D
 GENERAL INFORMATION:

APPLICANT: Sone, Toshio
 APPLICANT: Kume, Akimori
 APPLICANT: Daitiki, Kazuo
 APPLICANT: Iwama, Akiko
 APPLICANT: Kino, Kohsuke
 TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease
 FILE REFERENCE: SPO-103
 CURRENT APPLICATION NUMBER: US/09/142,524D
 CURRENT FILING DATE: 1998-09-09
 PRIOR APPLICATION NUMBER: PCT/JP97/00740
 PRIOR FILING DATE: 1997-03-10
 NUMBER OF SEQ ID NOS: 174
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 1
 LENGTH: 80
 TYPE: PRT
 ORGANISM: Cryptosporidia japonica

Initial Score = 13 Optimized Score = 13 Significance = 0.00
 Residue Identity = 100% Matches = 13 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 KSMKTVAFNFGPN
 MKXTVAFNFGPNRRVFIKRVSNVLIHGRIDIPASKNPHLQKNTIGTGRISLKLTSKIASRRVDGII
 X 10 X 20 30 40 50 60 70

AAAYONPASWK
 80

2. US-09-142-524D-57 (1-15)

US-09-142-524D-2 Sequence 2, Application US/09142524D

Sequence 2, Application US/09142524D
 GENERAL INFORMATION:

APPLICANT: Sone, Toshio
 APPLICANT: Kume, Akimori
 APPLICANT: Daitiki, Kazuo
 APPLICANT: Iwama, Akiko
 APPLICANT: Kino, Kohsuke
 TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease
 FILE REFERENCE: SPO-103
 CURRENT APPLICATION NUMBER: US/09/142,524D
 CURRENT FILING DATE: 1998-09-09
 PRIOR APPLICATION NUMBER: PCT/JP97/00740
 PRIOR FILING DATE: 1997-03-10
 NUMBER OF SEQ ID NOS: 174
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 2
 LENGTH: 105
 TYPE: PRT
 ORGANISM: Cryptosporidia japonica

Initial Score = 13 Optimized Score = 13 Significance = 0.00
 Residue Identity = 100% Matches = 13 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X 10 X

KSMKVTVAENQFGPN
 |||||
 MKVTVAENQFGPNRVFIKRVSNVLIHGRRIDIPASKNPHLOKNTIGTGRKWKNNRIWLOPAKLTGFTLM
 X 10 X 20 30 40 50 60 70
 GRRLKMPWYIAGYKTFDGRRVVDGIIAAYONPASMK
 80 90 100

3. US-09-142-524D-57 (1-15)
 US-09-142-524D-3 Sequence 3, Application US/09142524D

Sequence 3, Application US/09142524D
 GENERAL INFORMATION:

APPLICANT: Sone, Toshio
 APPLICANT: Kume, Akimori
 APPLICANT: Dairiki, Kazuo

APPLICANT: Iwama, Akiko
 APPLICANT: Kino, Kohsuke

TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Diseases
 FILE REFERENCE: SPO-103

CURRENT APPLICATION NUMBER: US/09/142,524D
 CURRENT FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: PCT/JP97/00740
 PRIOR FILING DATE: 1997-03-10

NUMBER OF SEQ ID NOS: 174
 SOFTWARE: PatentIn version 3.1

SEQ ID NO 3

LENGTH: 134

TYPE: PRT

ORGANISM: Cryptomeria japonica

Initial Score = 13 Optimized Score = 13 Significance = 0.00
 Residue Identity = 100% Matches = 13 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

X 10 X
 KSMKVTVAENQFGPN
 |||||
 MKVTVAENQFGPNRVFIKRVSNVLIHGRRIDIPASKNPHLOKNTIGTGRKWKNNRIWLOPAKLTGFTLM
 X 10 X 20 30 40 50 60 70
 GRRLKMPWYIAGYKTFDGRRVVDGIIAAYONPASMK
 80 90 100 110 120 130

=> fil reg

FILE 'REGISTRY' ENTERED AT 18:25:31 ON 23 APR 2003
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2003 American Chemical Society (ACS)

Property values tagged with IC are from the ZIC/VINITI data file
provided by InfoChem.

STRUCTURE FILE UPDATES: 22 APR 2003 HIGHEST RN 503805-80-9
DICTIONARY FILE UPDATES: 22 APR 2003 HIGHEST RN 503805-80-9

TSCA INFORMATION NOW CURRENT THROUGH JANUARY 6, 2003

Please note that search-term pricing does apply when
conducting SmartSELECT searches.

Crossover limits have been increased. See HELP CROSSOVER for details.

Experimental and calculated property data are now available. See HELP
PROPERTIES for more information. See STNote 27, Searching Properties
in the CAS Registry File, for complete details:
<http://www.cas.org/ONLINE/STN/STNOTES/stnotes27.pdf>

=> d sta que l16

L7 56 SEA FILE=REGISTRY ABB=ON PLU=ON (RPLWIIIFSGNMNKL)|(IFSGNMNKLKM
PMY)|(MNIKLKMPMYIAGYK)|(KMPMYIAGYKTFDGR)|(PCVFKRVSNVIIHG)|(KSMK
VTVAFNQFGPN)/SQSP
L15 24 SEA FILE=REGISTRY ABB=ON PLU=ON (VDGIIAAYQNPASWK)|(NNRIWLQFAK
LTGFT)|(LQFAKLTGFTLMGKG)|(GIDIFASKNFHLQKN)|(ASKNFHLQKNTIGTG)|(S
RAEVSYPVHVNGAKF)|(LSDISLKLTSKGIAS)/SQSP
L16 6 SEA FILE=REGISTRY ABB=ON PLU=ON L7 AND L15

=> d l16 sqide can tot

L16 ANSWER 1 OF 6 REGISTRY COPYRIGHT 2003 ACS
RN 502526-68-3 REGISTRY
CN 24: PN: JP2003079389 SEQID: 1 unclaimed protein (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE
SQL 105

PATENT ANNOTATIONS (PNTE):

Sequence |Patent
Source |Reference

=====+=====

Not Given	JP2003079389
	unclaimed
	SEQID 1

SEQ 1 MKVTVAFNQF GPNRRVFIKR VSNVIIHGRR IDIFASKNFH LQKNTIGTGR
=====

51 RWKNNRIWLQ FAKLTGFTLM GRRLKMPMYI AGYKTFDGRR VDGIIAAYQN
=====

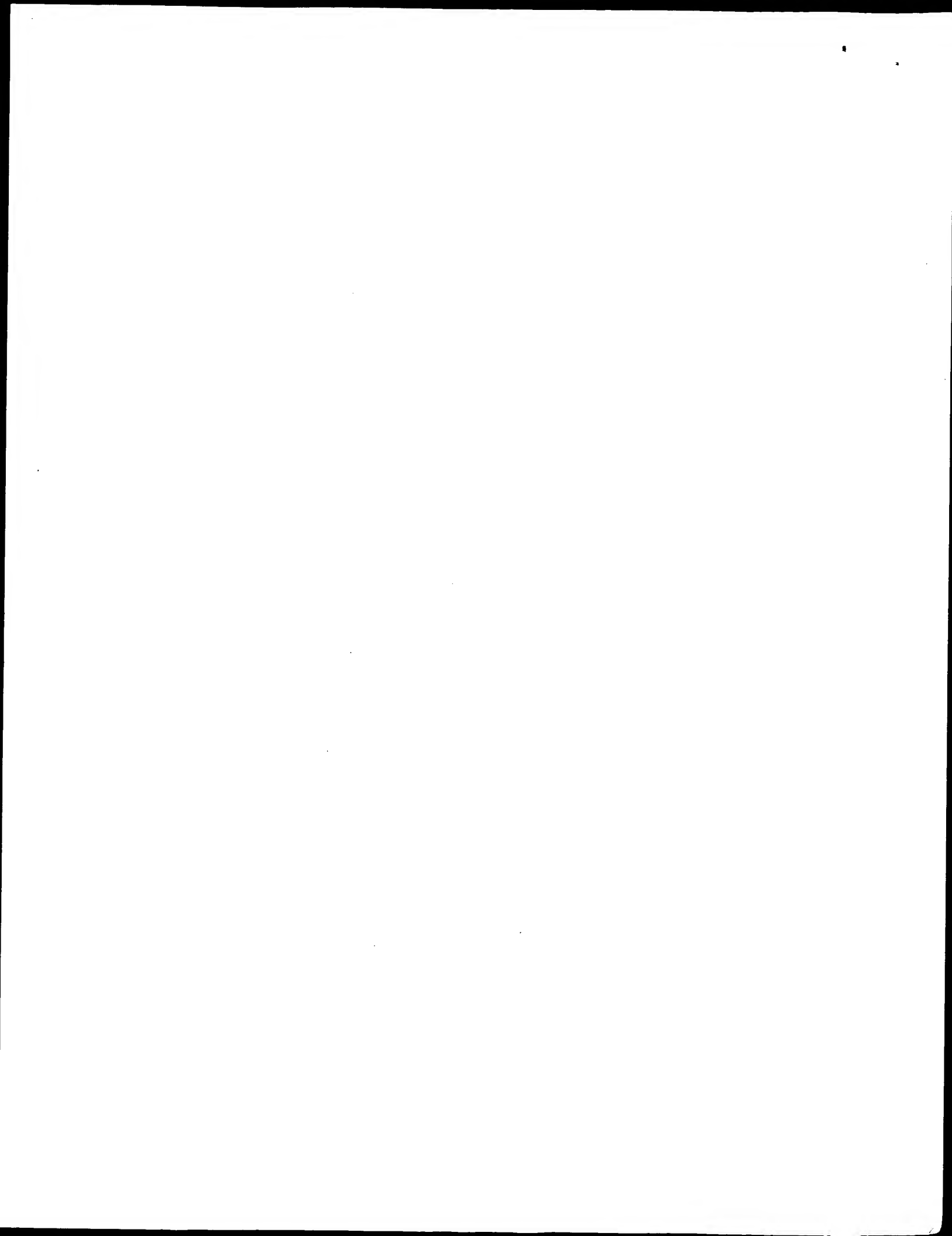
101 PASWK
=====

HITS AT: 35-49, 54-68, 75-89, 91-105

RELATED SEQUENCES AVAILABLE WITH SEQLINK

MF Unspecified
CI MAN
SR CA

Jan Delaval
Reference Librarian
Biology & Chemical Library
Call 1-800-703-3003-498
jan.delaval@uspis.gov



LC STN Files: CA, CAPLUS
1 REFERENCES IN FILE CA (1962 TO DATE)
1 REFERENCES IN FILE CAPLUS (1962 TO DATE)

REFERENCE 1: 138:253826

L16 ANSWER 2 OF 6 REGISTRY COPYRIGHT 2003 ACS
RN 502526-45-6 REGISTRY
CN 1: PN: JP2003079389 SEQID: 3 unclaimed protein (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE
SQL 134

PATENT ANNOTATIONS (PNTE):

Sequence	Patent
Source	Reference
=====+	
Not Given	JP2003079389
	unclaimed
	SEQID 3

SEQ 1 MKVTVAFNQF GPNRRVFIKR VSNVIIHGRR IDIFASKNFH LQKNTIGTGR
=====

51RWKNNRIWLQ FAKLTGFTLM GRRPLWIIFS GNMNIKLKMP MYIAGYKTFD
=====

101 GRRAEVSYVH VNGAKFIRRV DGIIAAYQNP ASWK
=====

HITS AT: 35-49, 54-68, 83-102, 120-134

RELATED SEQUENCES AVAILABLE WITH SEQLINK

MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS
1 REFERENCES IN FILE CA (1962 TO DATE)
1 REFERENCES IN FILE CAPLUS (1962 TO DATE)

REFERENCE 1: 138:253826

L16 ANSWER 3 OF 6 REGISTRY COPYRIGHT 2003 ACS
RN 501062-84-6 REGISTRY
CN Allergen (Cryptomeria japonica synthetic allergen multi-T cell epitope B)
(9CI) (CA INDEX NAME)
OTHER NAMES:
CN 3: PN: JP2003079387 SEQID: 3 claimed protein
FS PROTEIN SEQUENCE
SQL 134

PATENT ANNOTATIONS (PNTE):

Sequence	Patent
Source	Reference
=====+	
Not Given	JP2003079387
	claimed
	SEQID 3

SEQ 1 MKVTVAFNQF GPNRRVFIKR VSNVIIHGRR IDIFASKNFH LQKNTIGTGR
=====

51 RWKNNRIWLQ FAKLTGFTLM GRRPLWIIFS GNMNIKLKMP MYIAGYKTFD
=====

101 GRRAEVSYVH VNGAKFIRRV DGIIAAYQNP ASWK
=====



HITS AT: 35-49, 54-68, 83-102, 120-134

RELATED SEQUENCES AVAILABLE WITH SEQLINK

MF Unspecified

CI MAN

SR CA

LC STN Files: CA, CAPLUS

1 REFERENCES IN FILE CA (1962 TO DATE)

1 REFERENCES IN FILE CAPLUS (1962 TO DATE)

REFERENCE 1: 138:220362

L16 ANSWER 4 OF 6 REGISTRY COPYRIGHT 2003 ACS

RN 501062-83-5 REGISTRY

CN Allergen (Cryptomeria japonica synthetic allergen multi-T cell epitope A)
(9CI) (CA INDEX NAME)

OTHER NAMES:

CN 1: PN: JP2003079387 SEQID: 1 claimed protein

FS PROTEIN SEQUENCE

SQL 105

PATENT ANNOTATIONS (PNTE):

Sequence |Patent

Source |Reference

====+=====

Not Given|JP2003079387

|claimed

|SEQID 1

SEQ 1 MKVTVAFNQF GPNRRVFIKR VSNVIIHGRR IDIFASKNFH LQKNTIGTGR

51 RWKNNRIWLQ FAKLTGFTLM GRRLKMPMYI AGYKTFDGRR VDGIIAAYQN

101 PASWK

HITS AT: 35-49, 54-68, 75-89, 91-105

RELATED SEQUENCES AVAILABLE WITH SEQLINK

MF Unspecified

CI MAN

SR CA

LC STN Files: CA, CAPLUS

1 REFERENCES IN FILE CA (1962 TO DATE)

1 REFERENCES IN FILE CAPLUS (1962 TO DATE)

REFERENCE 1: 138:220362

L16 ANSWER 5 OF 6 REGISTRY COPYRIGHT 2003 ACS

RN 196718-37-3 REGISTRY

CN Multiepitope allergen (synthetic Cryptomeria japonica allergen Cry j 1/Cry
j 2-derived 134-amino acid isoform) (9CI) (CA INDEX NAME)

FS PROTEIN SEQUENCE

SQL 134

SEQ 1 MKVTVAFNQF GPNRRVFIKR VSNVIIHGRR IDIFASKNFH LQKNTIGTGR

51 RWKNNRIWLQ FAKLTGFTLM GRRPLWIIFS GNMNIKLKMP MYIAGYKTFD

101 GRRAEVSIVH VNGAKFIRRV DGIIAAYQNP ASWK

HITS AT: 35-49, 54-68, 83-102, 120-134

RELATED SEQUENCES AVAILABLE WITH SEQLINK

MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS
1 REFERENCES IN FILE CA (1962 TO DATE)
1 REFERENCES IN FILE CAPLUS (1962 TO DATE)

REFERENCE 1: 127:277192

L16 ANSWER 6 OF 6 REGISTRY COPYRIGHT 2003 ACS
RN 196718-36-2 REGISTRY
CN Multiepitope allergen (synthetic Cryptomeria japonica allergen Cry j 1/Cry
j 2-derived 105-amino acid isoform) (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE
SQL 105

SEQ 1 MKVTVAFNQF GPNRRVFIKR VSNVIIHGRR IDIFASKNFH LQKNTIGTGR
=====
51 RWKNNRIWLQ FAKLTGFTLM GRRLKMPMYI AGYKTFDGRR VDGIIAAYQN
=====
101 PASWK
=====

HITS AT: 35-49, 54-68, 75-89, 91-105

RELATED SEQUENCES AVAILABLE WITH SEQLINK

MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS
1 REFERENCES IN FILE CA (1962 TO DATE)
1 REFERENCES IN FILE CAPLUS (1962 TO DATE)

REFERENCE 1: 127:277192

=> d his l16-

(FILE 'REGISTRY' ENTERED AT 18:15:33 ON 23 APR 2003)
L16 6 S L7 AND L15
SAV L7 DIBRINO1/A
SAV L15 DIBRINO2/A
SAV L16 DIBRINO3/A

FILE 'HCAOLD' ENTERED AT 18:22:48 ON 23 APR 2003
L17 0 S L16

FILE 'HCAPLUS' ENTERED AT 18:22:56 ON 23 APR 2003
L18 3 S L16
SEL RN
DEL SEL
SEL RN 3

FILE 'REGISTRY' ENTERED AT 18:23:36 ON 23 APR 2003
L19 5 S E1-E5
L20 2 S L19 AND L16
L21 3 S L19 NOT L20
L22 1 S L21 AND L7,L15
L23 2 S L21 NOT L22

FILE 'USPATFULL, USPAT2' ENTERED AT 18:25:15 ON 23 APR 2003
L24 0 S L16

FILE 'REGISTRY' ENTERED AT 18:25:31 ON 23 APR 2003

=> fil hcaplus

FILE 'HCAPLUS' ENTERED AT 18:25:56 ON 23 APR 2003

USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.

PLEASE SEE "HELP USAGETERMS" FOR DETAILS.

COPYRIGHT (C) 2003 AMERICAN CHEMICAL SOCIETY (ACS)

Copyright of the articles to which records in this database refer is held by the publishers listed in the PUBLISHER (PB) field (available for records published or updated in Chemical Abstracts after December 26, 1996), unless otherwise indicated in the original publications. The CA Lexicon is the copyrighted intellectual property of the American Chemical Society and is provided to assist you in searching databases on STN. Any dissemination, distribution, copying, or storing of this information, without the prior written consent of CAS, is strictly prohibited.

FILE COVERS 1907 - 23 Apr 2003 VOL 138 ISS 17

FILE LAST UPDATED: 22 Apr 2003 (20030422/ED)

This file contains CAS Registry Numbers for easy and accurate substance identification.

=> d 118 all tot

L18 ANSWER 1 OF 3 HCAPLUS COPYRIGHT 2003 ACS

AN 2003:210113 HCAPLUS

DN 138:253826

TI Purification of recombinant multi-T cell epitope from inclusion body

IN Miyahara, Michinori; Mori, Takeshi; Yokoyama, Minehiko; Kamiyo, Masayuki; Yamazaki, Tetsuya; Yamashita, Akio; Osamu, Kosuke

PA Meiji Milk Products, Co., Ltd., Japan

SO Jpn. Kokai Tokkyo Koho, 16 pp.

CODEN: JKXXAF

DT Patent

LA Japanese

IC ICM C12P021-02

ICS C12N015-09

CC 16-1 (Fermentation and Bioindustrial Chemistry)

Section cross-reference(s): 1, 3, 15

FAN.CNT 1

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	JP 2003079389	A2	20030318	JP 2002-189240	20020628
PRAI	JP 2001-196579	A	20010628		

AB The recombinant multi-T cell epitope polypeptide is purified from inclusion body of Escherichia coli by solubilization with guanidine/urea; and subjection to copper-chelated chromatog., cation exchange chromatog., and reverse-phase chromatog. Purifn. of recombinant multi-T cell epitope polypeptide A comprising 6 epitopes of Cry J1 and j2, the main allergen protein of Cryptomeria japonica, by construction of plasmid pQTF7.DELTA.cr, transformation of E. coli, etc. was shown.

ST sugi pollen allergen protein epitope chromatog purifn; Escherichia recombinant sugi pollen allergen protein epitope

IT Allergens

RL: BPN (Biosynthetic preparation); PUR (Purification or recovery); BIOL (Biological study); PREP (Preparation)

(Cry j I (Cryptomeria japonica, I); purifn. of recombinant multi-T cell epitope from inclusion body of Escherichia coli)

IT Allergens

RL: BPN (Biosynthetic preparation); PUR (Purification or recovery); BIOL (Biological study); PREP (Preparation)
 (Cry j II; purifn. of recombinant multi-T cell epitope from inclusion body of Escherichia coli)

IT Affinity chromatography
 (copper-chelated; purifn. of recombinant multi-T cell epitope from inclusion body of Escherichia coli)

IT Plasmids
 (pQTF7.DELTA.cr; purifn. of recombinant multi-T cell epitope from inclusion body of Escherichia coli)

IT Cryptomeria japonica
 (pollen; purifn. of recombinant multi-T cell epitope from inclusion body of Escherichia coli)

IT Cation exchange chromatography
 DNA sequences

Epitopes

Escherichia coli

Fermentation

Inclusion bodies

Pollen

Protein sequences

Reversed phase chromatography

(purifn. of recombinant multi-T cell epitope from inclusion body of Escherichia coli)

IT Immunization
 (vaccination; purifn. of recombinant multi-T cell epitope from inclusion body of Escherichia coli)

IT 57-13-6, Urea, biological studies 113-00-8, Guanidine
 RL: BUU (Biological use, unclassified); BIOL (Biological study); USES (Uses)

(purifn. of recombinant multi-T cell epitope from inclusion body of Escherichia coli)

IT 502526-46-7 502526-47-8 502526-48-9 502526-49-0 502526-50-3
 502526-51-4 502526-52-5 502526-53-6 502526-54-7 502526-55-8
 502526-56-9 502526-57-0 502526-58-1 502526-59-2 502526-60-5
 502526-61-6 502526-62-7 502526-63-8 502526-64-9 502526-65-0
 502526-66-1 502526-67-2 502526-69-4

RL: PRP (Properties)

(unclaimed nucleotide sequence; purifn. of recombinant multi-T cell epitope from inclusion body)

IT 502526-45-6 502526-68-3

RL: PRP (Properties)

(unclaimed protein sequence; purifn. of recombinant multi-T cell epitope from inclusion body)

L18 ANSWER 2 OF 3 HCAPLUS COPYRIGHT 2003 ACS

AN 2003:210110 HCAPLUS

DN 138:220362

TI Expression of recombinant Japanese cedar allergen multi-T cell epitope in Escherichia coli inclusion body

IN Miyahara, Michinori; Mori, Takeshi; Yokoyama, Minehiko; Kamijo, Masayuki; Yamazaki, Tetsuya; Yamashita, Akio; Kino, Kouuke

PA Meiji Milk Products, Co., Ltd., Japan

SO Jpn. Kokai Tokkyo Koho, 19 pp.

CODEN: JKXXAF

DT Patent

LA Japanese

IC ICM C12N015-09

ICS A61K035-74; A61K048-00; A61P037-00

CC 15-2 (Immunochemistry)

Section cross-reference(s): 3, 16

FAN.CNT 1

PATENT NO.

KIND DATE

APPLICATION NO.

DATE

----- X -----

PI JP 2003079387 A2 20030318 JP 2002-189243 20020628

PRAI JP 2001-196598 A 20010628

AB Recombinant expression of multi-T cell epitopes comprising 6 or 7 epitopes of major allergens of Japanese cedar (*C. japonica*) pollen, Cry j 1 and Cry j2, in *E. coli* as inclusion bodies, is disclosed. A construct contg. trp operon promoter/operator region, Shine-Dalgarno sequence, multi-T cell epitope coding sequence, and bacteriophage .lambda. terminator to, is used.

ST sugi pollen allergen epitope recombinant expression *Escherichia* inclusion body; T cell epitope *Cryptomeria* pollen allergen *Escherichia* inclusion body; Japanese cedar allergen epitope *Escherichia* inclusion body expression

IT Allergens
RL: BPN (Biosynthetic preparation); PRP (Properties); BIOL (Biological study); PREP (Preparation)
(Cry j I (*Cryptomeria japonica*, I); expression of recombinant Japanese cedar allergen multi-T cell epitope in *Escherichia coli* inclusion body)

IT Allergens
RL: BPN (Biosynthetic preparation); PRP (Properties); BIOL (Biological study); PREP (Preparation)
(Cry j II; expression of recombinant Japanese cedar allergen multi-T cell epitope in *Escherichia coli* inclusion body)

IT *Cryptomeria japonica*
Epitopes
Escherichia coli
Inclusion bodies
Pollen
Protein sequences
T cell (lymphocyte)
cDNA sequences
(expression of recombinant Japanese cedar allergen multi-T cell epitope in *Escherichia coli* inclusion body)

IT Genetic element
RL: BUU (Biological use, unclassified); BIOL (Biological study); USES (Uses)
(operator, trp operon; expression of recombinant Japanese cedar allergen multi-T cell epitope in *Escherichia coli* inclusion body)

IT Genetic element
RL: BUU (Biological use, unclassified); BIOL (Biological study); USES (Uses)
(ribosome-binding site; expression of recombinant Japanese cedar allergen multi-T cell epitope in *Escherichia coli* inclusion body)

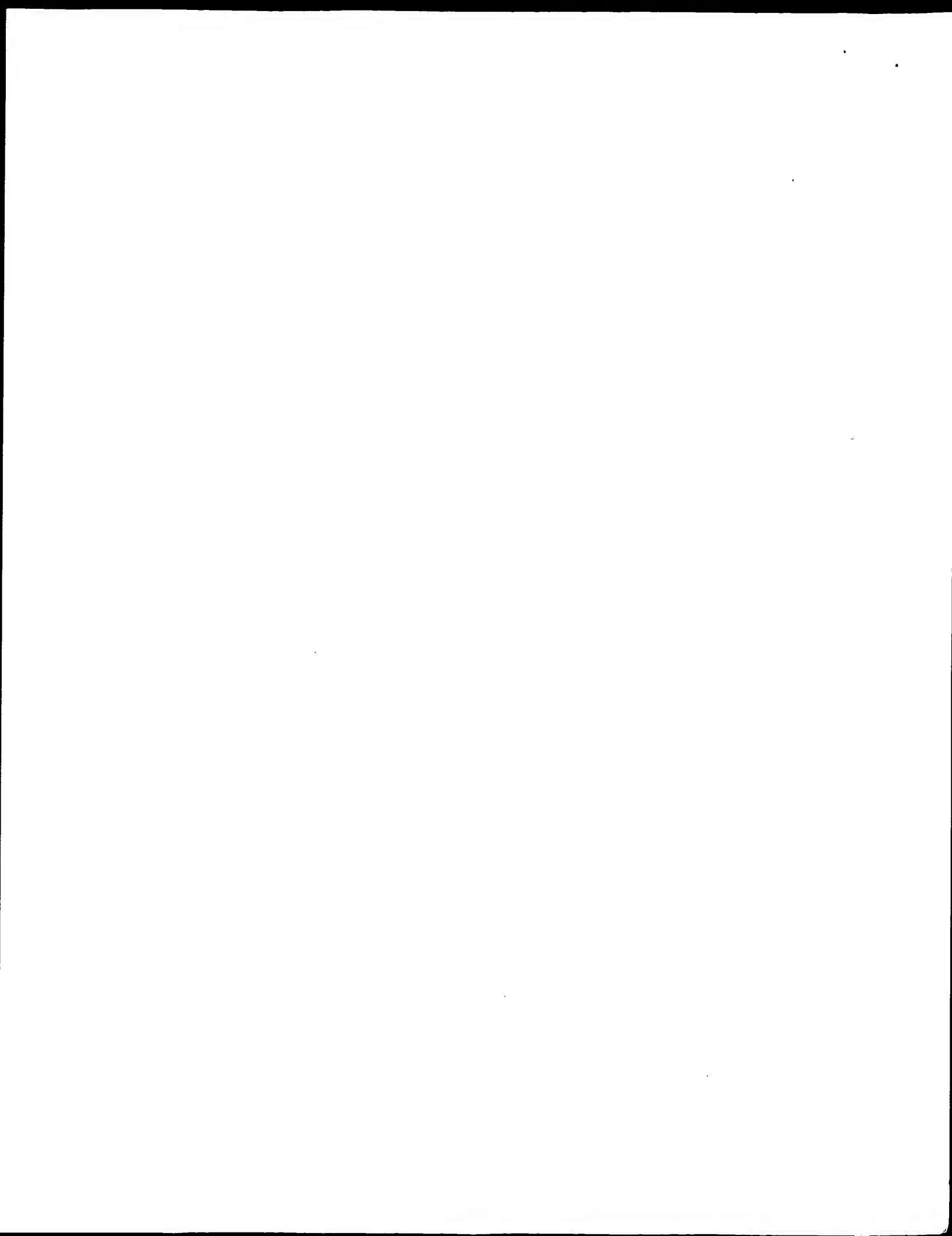
IT Coliphage .lambda.
(terminator of; expression of recombinant Japanese cedar allergen multi-T cell epitope in *Escherichia coli* inclusion body)

IT Genetic element
RL: BUU (Biological use, unclassified); BIOL (Biological study); USES (Uses)
(terminator, bacteriophage .lambda.; expression of recombinant Japanese cedar allergen multi-T cell epitope in *Escherichia coli* inclusion body)

IT Promoter (genetic element)
RL: BUU (Biological use, unclassified); BIOL (Biological study); USES (Uses)
(trp operon; expression of recombinant Japanese cedar allergen multi-T cell epitope in *Escherichia coli* inclusion body)

IT Operon
(trp; expression of recombinant Japanese cedar allergen multi-T cell epitope in *Escherichia coli* inclusion body)

IT **501062-83-5P**
RL: BPN (Biosynthetic preparation); PRP (Properties); BIOL (Biological study); PREP (Preparation)
(amino acid sequence; expression of recombinant Japanese cedar allergen



multi-T cell epitope in Escherichia coli inclusion body)

IT 501062-84-6
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; expression of recombinant Japanese cedar allergen multi-T cell epitope in Escherichia coli inclusion body)

IT 501062-85-7
RL: BSU (Biological study, unclassified); BUU (Biological use, unclassified); PRP (Properties); BIOL (Biological study); USES (Uses)
(nucleotide sequence; expression of recombinant Japanese cedar allergen multi-T cell epitope in Escherichia coli inclusion body)

IT 501062-86-8
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(nucleotide sequence; expression of recombinant Japanese cedar allergen multi-T cell epitope in Escherichia coli inclusion body)

IT 501070-61-7 501070-62-8 501070-63-9 501070-64-0 501070-65-1
501070-66-2 501070-67-3 501070-68-4 501070-69-5 501070-70-8
501070-71-9 501070-72-0 501070-73-1 501070-74-2 501070-75-3
501070-76-4 501070-77-5 501070-78-6 501070-79-7 501070-80-0
501070-81-1 501070-82-2
RL: PRP (Properties)
(unclaimed nucleotide sequence; expression of recombinant Japanese cedar allergen multi-T cell epitope in Escherichia coli inclusion body)

L18 ANSWER 3 OF 3 HCAPLUS COPYRIGHT 2003 ACS

AN 1997:617989 HCAPLUS

DN 127:277192

TI Peptide-base immunotherapeutic agent for allergic diseases

IN Sone, Toshio; Kume, Akinori; Dairiki, Kazuo; Iwama, Akiko; Kino, Kohsuke

PA Meiji Milk Products Co., Ltd., Japan; Sone, Toshio; Kume, Akinori;

Dairiki, Kazuo; Iwama, Akiko; Kino, Kohsuke

SO PCT Int. Appl., 58 pp.

CODEN: PIXXD2

DT Patent

LA Japanese

IC ICM A61K039-35

ICS C07K014-725

CC 15-2 (Immunochemistry)

FAN.CNT 1

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9732600	A1	19970912	WO 1997-JP740	19970310
W: CA, CN, JP, KR, US				
RW: AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
CA 2248937	AA	19970912	CA 1997-2248937	19970310
CN 1218412	A	19990602	CN 1997-194541	19970310
EP 923940	A1	19990623	EP 1997-906863	19970310
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				

PRAI JP 1996-80702 19960310

WO 1997-JP740 19970310

AB A peptide-base immunotherapeutic agent contg. an efficacious amt. of a monomol. multi-epitope peptide prepd. by bonding T-cell epitope regions derived from different allergen mols. with each other, e.g. Cry j 1/Cry i 2 or Cry j 1/Cha 0 1. It is efficacious for the prevention and therapy of a wide variety of allergic diseases.

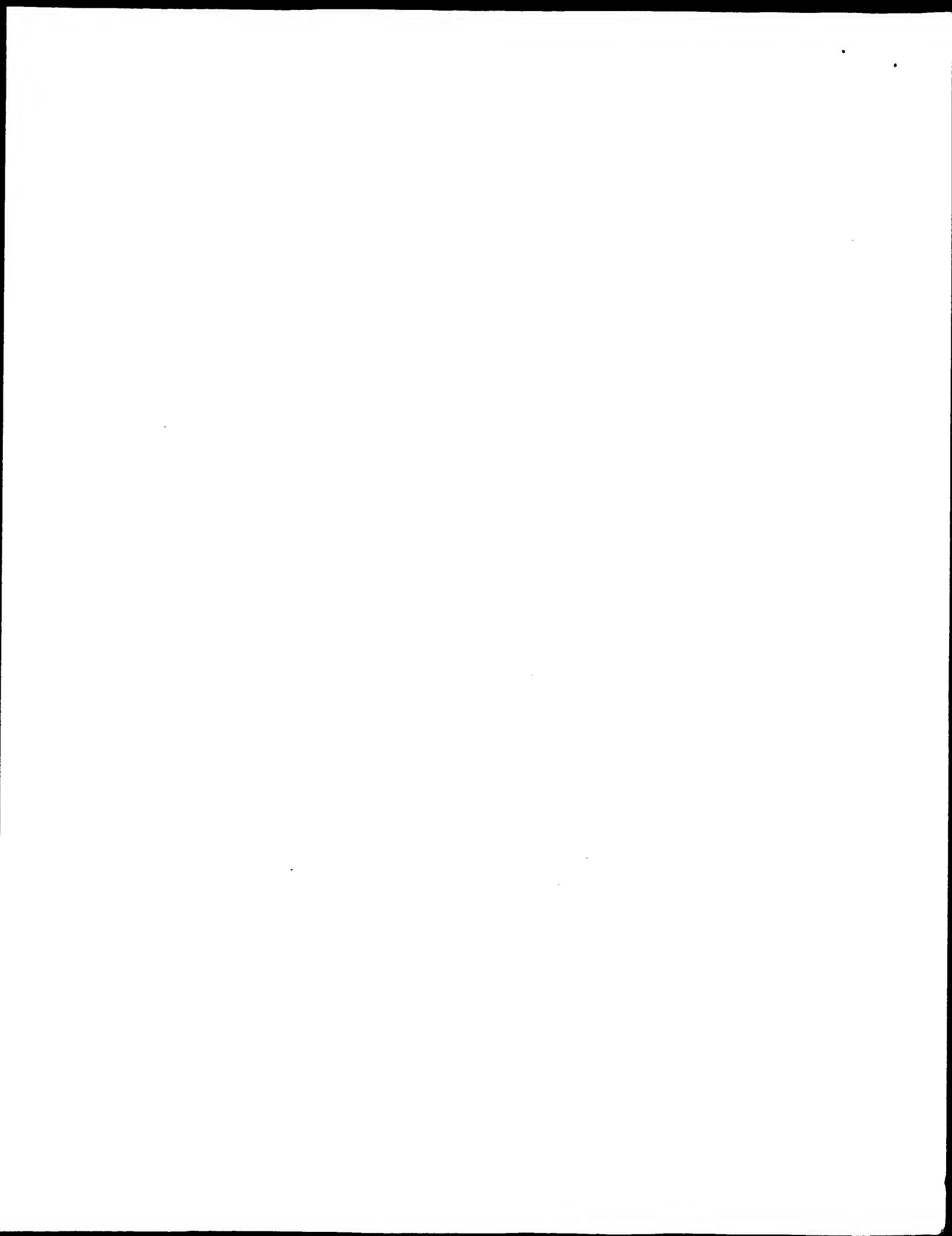
ST allergen multiple epitope immunotherapy allergic disease

IT Allergens

RL: BSU (Biological study, unclassified); BIOL (Biological study)

(Cha 0 1 multi-epitope; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)

- IT Allergens
RL: BSU (Biological study, unclassified); BIOL (Biological study)
(Cry j I (Cryptomeria japonica, I), multi-epitope; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT Allergens
RL: BSU (Biological study, unclassified); BIOL (Biological study)
(Cry j II, multi-epitope; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT Immunoglobulins
RL: BSU (Biological study, unclassified); BIOL (Biological study)
(E; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT Histocompatibility antigens
RL: BPR (Biological process); BSU (Biological study, unclassified); BIOL (Biological study); PROC (Process)
(HLA-DP, DPB1*0501 and DPA1*0101 and DPB1*0201; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT Histocompatibility antigens
RL: BPR (Biological process); BSU (Biological study, unclassified); BIOL (Biological study); PROC (Process)
(HLA-DQ, DQB1*0602 and DQA1*0102; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT Histocompatibility antigens
RL: BPR (Biological process); BSU (Biological study, unclassified); BIOL (Biological study); PROC (Process)
(HLA-DR, DRB5*0101 and DRB4*0101; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT Histocompatibility antigens
RL: BPR (Biological process); BSU (Biological study, unclassified); BIOL (Biological study); PROC (Process)
(HLA; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT Epitopes
RL: BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)
(T cell; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT Antigen presentation
(cells; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT T cell (lymphocyte)
(epitopes; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT Cryptomeria japonica
Immunotherapy
Lymphocyte
Protein sequences
(multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT Allergens
RL: BPN (Biosynthetic preparation); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)
(multi-epitope; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT 196718-35-1P 196718-36-2P 196718-37-3P
RL: BPN (Biosynthetic preparation); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)
(amino acid sequence; multi-epitope peptides of Cry j 1 and/or Cry j 2



and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)

IT 196510-89-1P 196521-96-7P
RL: BPN (Biosynthetic preparation); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)
(multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)

=> fil reg

FILE 'REGISTRY' ENTERED AT 18:26:09 ON 23 APR 2003
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2003 American Chemical Society (ACS)

Property values tagged with IC are from the ZIC/VINITI data file provided by InfoChem.

STRUCTURE FILE UPDATES: 22 APR 2003 HIGHEST RN 503805-80-9
DICTIONARY FILE UPDATES: 22 APR 2003 HIGHEST RN 503805-80-9

TSCA INFORMATION NOW CURRENT THROUGH JANUARY 6, 2003

Please note that search-term pricing does apply when conducting SmartSELECT searches.

Crossover limits have been increased. See HELP CROSSOVER for details.

Experimental and calculated property data are now available. See HELP PROPERTIES for more information. See STNote 27, Searching Properties in the CAS Registry File, for complete details:
<http://www.cas.org/ONLINE/STN/STNOTES/stnotes27.pdf>

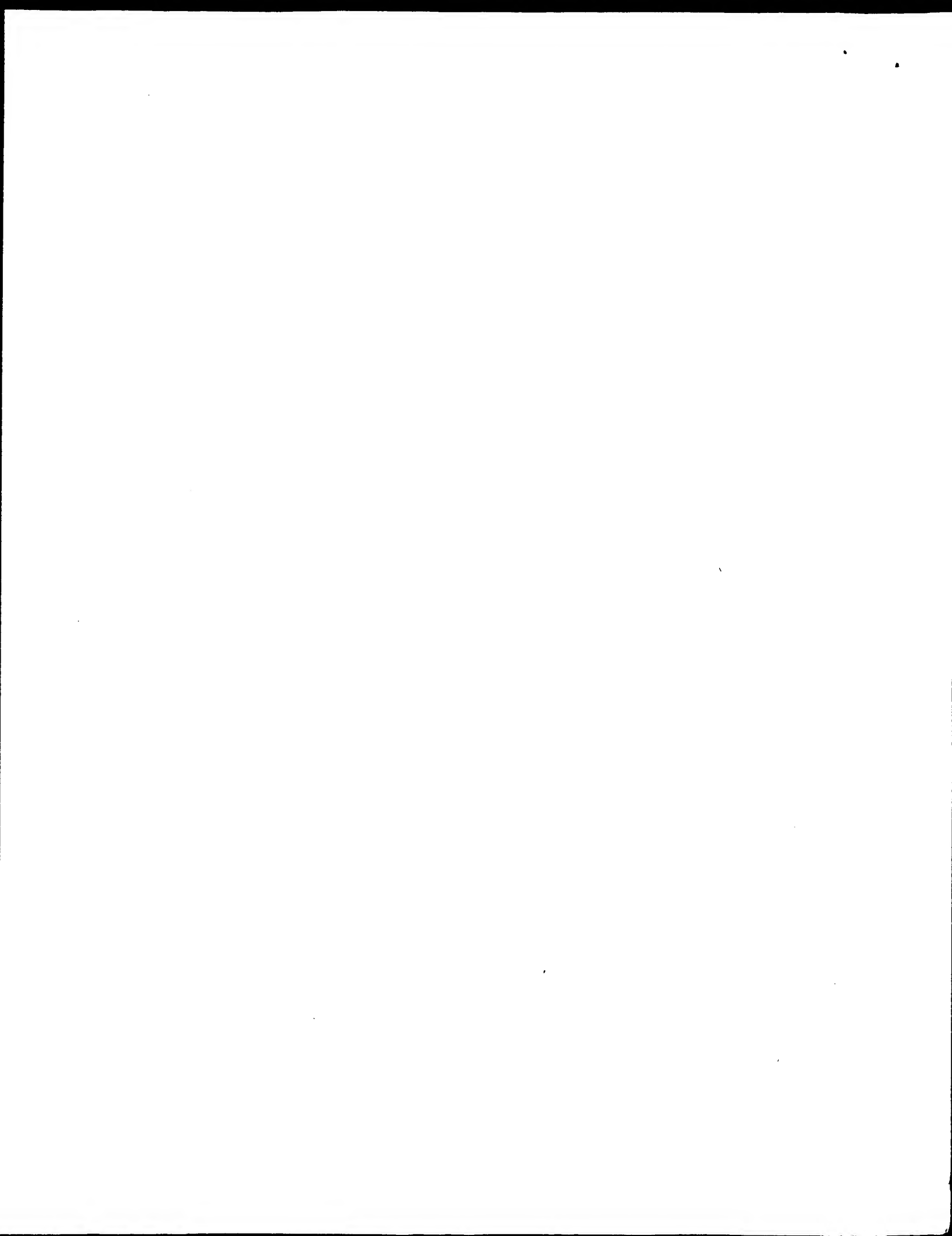
=> d sqide can tot 121

L21 ANSWER 1 OF 3 REGISTRY COPYRIGHT 2003 ACS
RN 196718-35-1 REGISTRY
CN Multi-epitope allergen (synthetic *Cryptomeria japonica* allergen Cry j 1/Cry j 2-derived 80-amino acid isoform) (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE
SQL 80

SEQ 1 MKVTVAFNQF GPNRRVFIKR VSNVIIHGRR IDIFASKNFH LQKNTIGTGR
51 RISLKLTS GK IASRRVDGII AAYQNPASWK
MF Unspecified
CI MAN
SR CA
LC STN Files: CA, CAPLUS
1 REFERENCES IN FILE CA (1962 TO DATE)
1 REFERENCES IN FILE CAPLUS (1962 TO DATE)

REFERENCE 1: 127:277192

L21 ANSWER 2 OF 3 REGISTRY COPYRIGHT 2003 ACS
RN 196521-96-7 REGISTRY
CN L-Isoleucine, L-isoleucyl-L-phenylalanyl-L-seryl-L-lysyl-L-asparaginyl-L-leucyl-L-asparaginyl-L-isoleucyl-L-lysyl-L-leucyl-L-asparaginyl-L-methionyl-L-prolyl-L-leucyl-L-tyrosyl-L-isoleucyl-L-alanylglycyl-L-asparaginyl-L-lysyl-L-arginyl-L-arginyl-L-phenylalanyl-L-isoleucyl-L-lysyl-L-arginyl-L-valyl-L-seryl-L-asparaginyl-L-valyl- (9CI) (CA INDEX NAME)
FS PROTEIN SEQUENCE
SQL 31



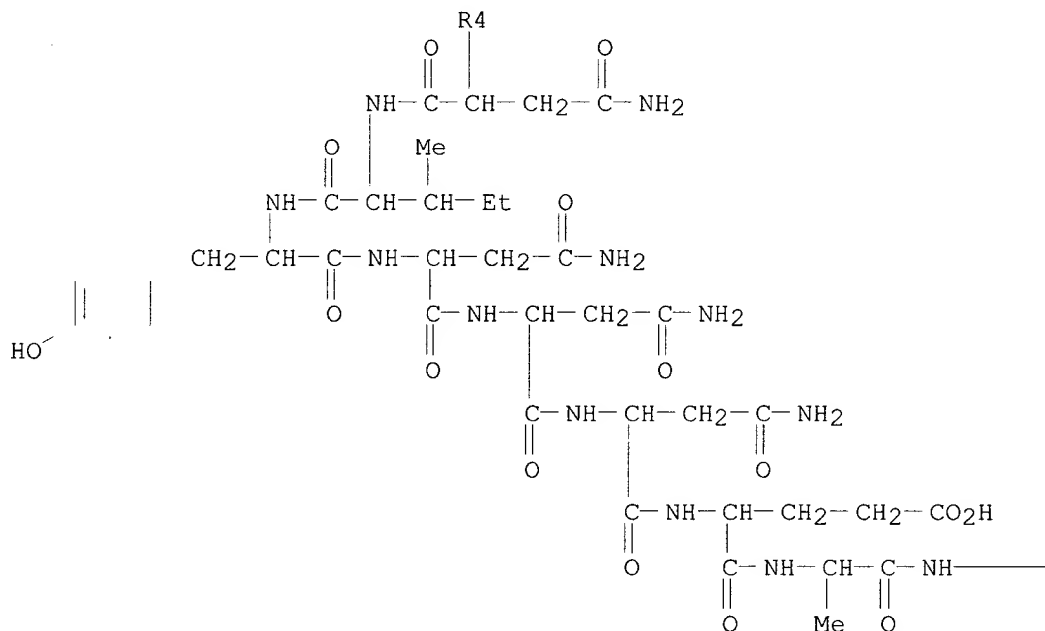
```
SEQ      1 IFSKNLNIKL NMPLYIAGNK RRFIKRVSNV I
MF C168 H283 N49 O40 S
CI MAN
SR CA
LC STN Files:   CA, CAPLUS
                1 REFERENCES IN FILE CA (1962 TO DATE)
                1 REFERENCES IN FILE CAPLUS (1962 TO DATE)
```

REFERENCE 1: 127:277192

```
L21  ANSWER 3 OF 3  REGISTRY  COPYRIGHT 2003  ACS
RN   196510-89-1  REGISTRY
CN   L-Isoleucine, L-seryl-L-serylglycyl-L-lysyl-L-asparaginyl-L-.alpha.-
      glutamylglycyl-L-threonyl-L-asparaginyl-L-isoleucyl-L-tyrosyl-L-
      asparaginyl-L-asparaginyl-L-asparaginyl-L-.alpha.-glutamyl-L-alanyl-L-
      phenylalanyl-L-lysyl-L-valyl-L-.alpha.-glutamyl-L-arginyl-L-arginyl-L-
      phenylalanyl-L-isoleucyl-L-lysyl-L-arginyl-L-valyl-L-seryl-L-asparaginyl-L-
      valyl- (9CI) (CA INDEX NAME)
FS   PROTEIN SEQUENCE; STEREOSEARCH
SQL  31
```

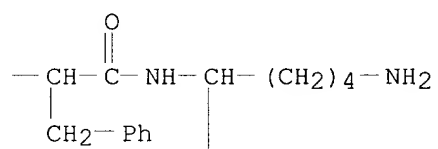
```
SEQ      1 SSGKNEGTNI YNNNEAFKVE RRFIKRVSNV I
MF      C155 H251 N49 O49
SR      CA
LC      STN Files:  CA, CAPLUS
```

PAGE 1-A

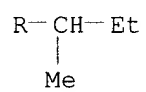




PAGE 1-B

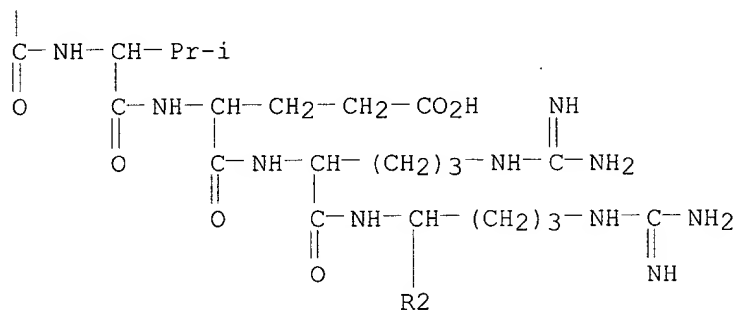


PAGE 2-A

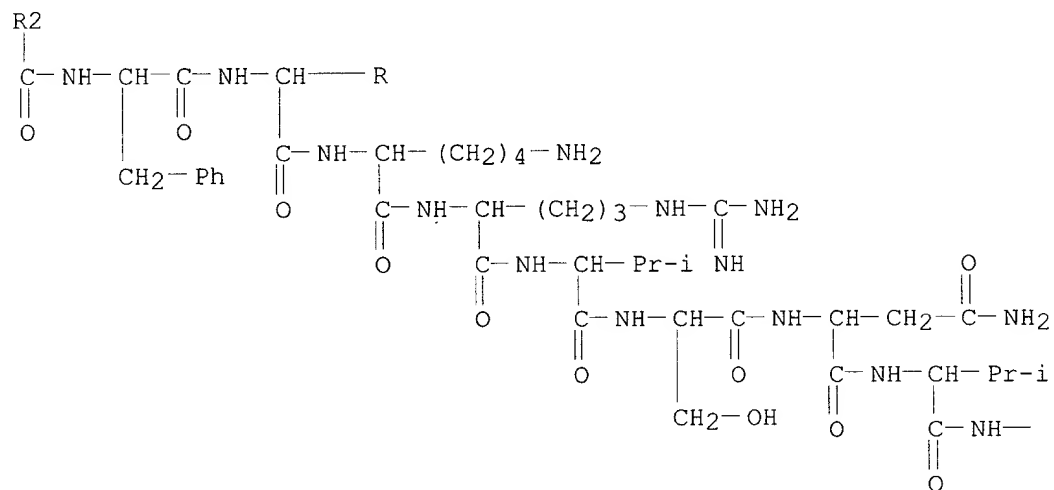




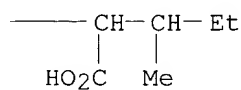
PAGE 2-B

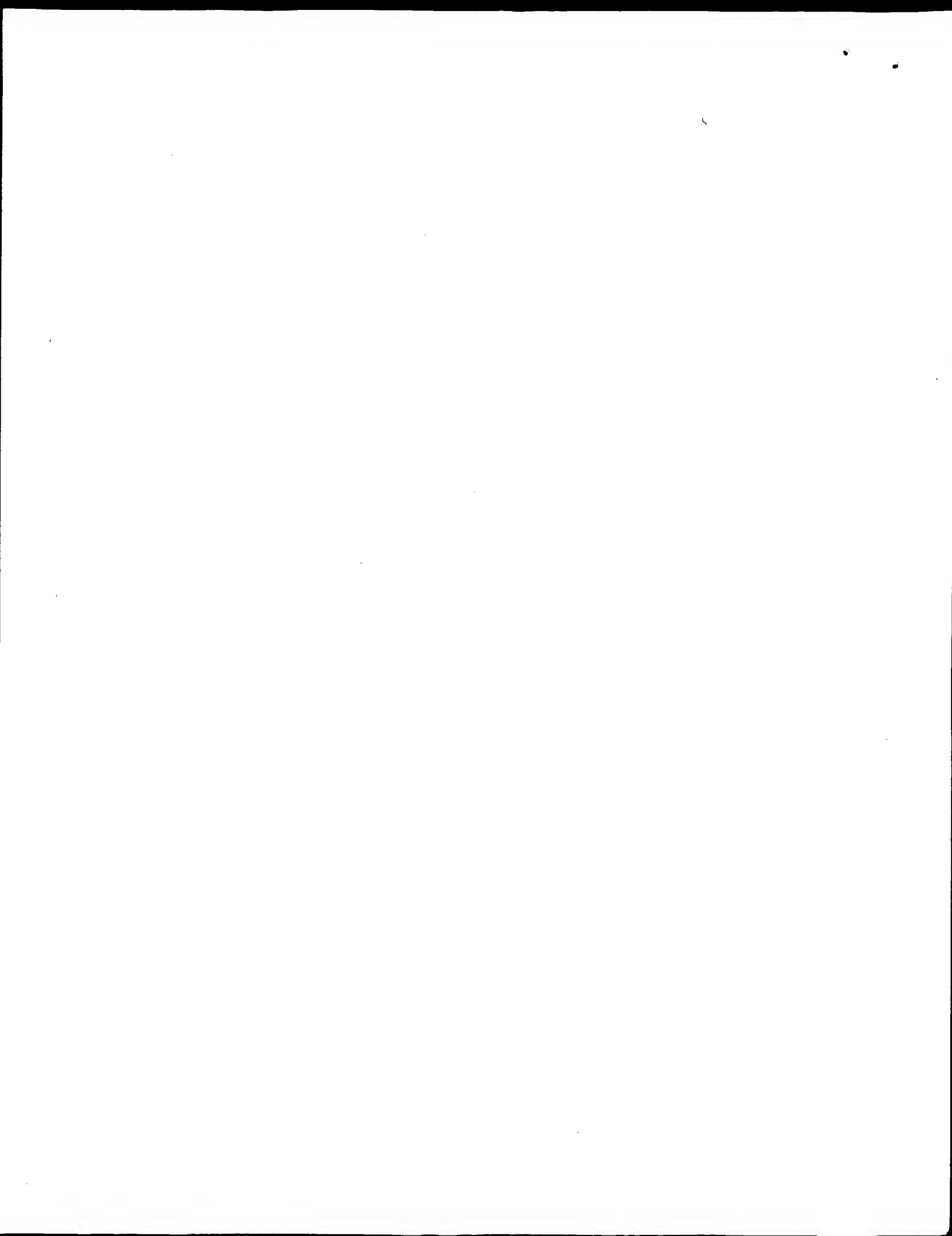


PAGE 3-A

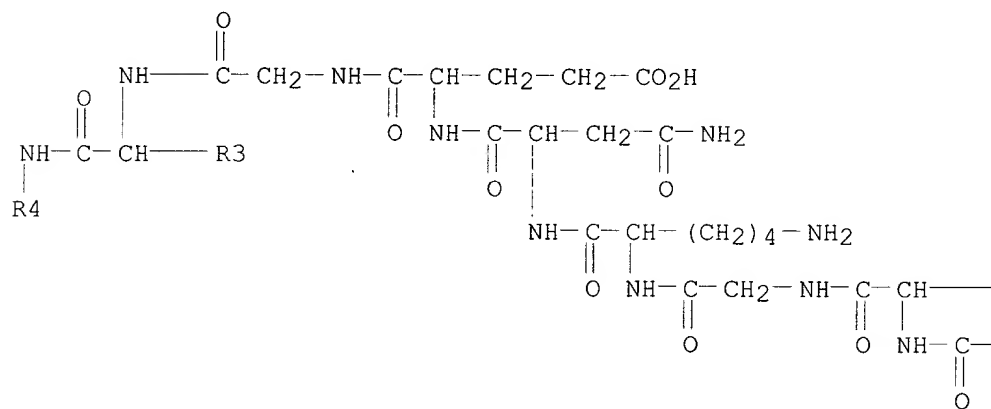
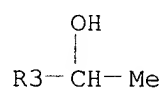


PAGE 3-B

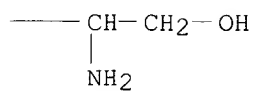
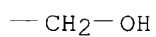




PAGE 4-A



PAGE 4-B



1 REFERENCES IN FILE CA (1962 TO DATE)
1 REFERENCES IN FILE CAPLUS (1962 TO DATE)

REFERENCE 1: 127:277192

